Scientific Research Associates, inc.

50 Nye Road, P.O. Box 1058
Tel: (203) 659-0333

Glastonbury, Connecticut 06033
Fax: (203) 633-0676

REPRINTS

THE PHYSICS AND OPERATION OF
ULTRA-SUBMICRON LENGTH SEMICONDUCTOR DEVICES

DTIC
ELECTE
JUN 0 9 1994
S F D

DTIC QUALITY INSPECTED 2

94-17517

94 6 8 123

DISCLAIMER NOTICE
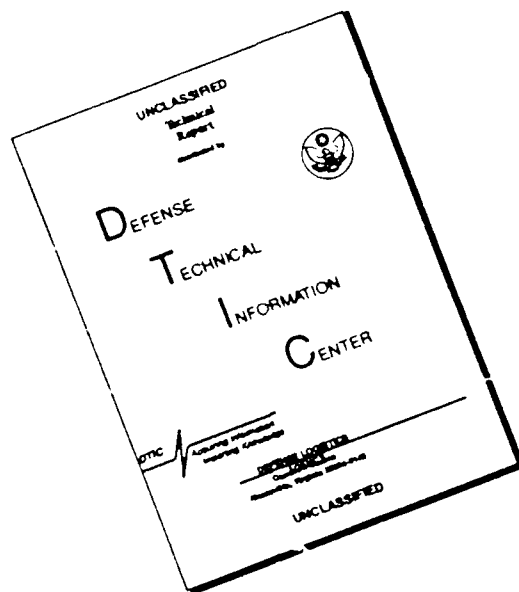
THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

ELECTRON DIFFRACTION THROUGH
AN APERTURE IN A POTENTIAL WALL

M. Cahay, J. P. Kreskovsky and H. L. Grubin

Scientific Research Associates, Inc., P. O. Box 1058
Glastonbury, CT 06033

Abstract

We study the diffraction of a two-dimensional Gaussian wavepacket through a rectangular aperture in a finite potential wall (one slit experiment). For wavepacket with incident wavevector k_0 satisfying the diffraction condition, $k_0 = 2\pi/w$, w being the slit width, the near field (Fresnel-like) diffraction pattern behind the slit can be clearly seen for small time duration (< 0.2ps). At later time steps, the diffracted beam is fragmented into lobes (perpendicular to the direction of incidence of the wavepacket) as a result of the multiple reflections of the wavepacket inside the slit (assumed to be of finite thickness). At later time, no far-field Fraunhoffer diffraction pattern is observed in our numerical simulations.

Introduction

With the advent of sophisticated growth techniques such as Molecular Beam Epitaxy and Metal Organic Chemical Vapor Deposition, there has been an increased theoretical interest in various quantum mechanical tunneling problems including: (1) the resonant tunneling of electrons through double barrier heterostructures, a problem of primary importance in asserting the switching time of resonant tunneling device (RTD) (Hauge and Stovneng, 1989; Goldberg and coworkers, 1967), (2) the electron propagation through narrow ballistic constriction (defined by a split-gate) in the two-dimensional electron gas of a $GaAs-Al_xGa_{1-x}As$ heterostructure. The latter followed the recent experimental discovery (Van Wees and coworkers, 1988) that the conductance of such constrictions increases in a sequence of steps of height $2e^2/h$ (at sufficiently low temperature). More recently, the possibility of using narrow split-gate for transistor applications has been suggested by Kriman and coworkers (1988) in their newly proposed QUADFET.

Both RTD's and QUADFET's have potential high-speed device applications within the terahertz regime (Hauge and Stovneng, 1989; Bandyopadhyay and coworkers, 1989). On one hand, the fast switching response of RTD's has been widely investigated using purely one dimensional analysis of wavepacket propagation through double barrier heterostructures (Hauge and Stovneng, 1989; Goldberg and coworkers, 1967). In practice however, electrons are injected from 3D contacts and tunneling through the quantum well of the RTD is characterized by two-dimensional dynamics. On the other hand, in view of their potentiality for high-speed applications (such as in the QUADFET), there is now a call for a transient analysis of quantum transport through narrow ballistic constrictions to supplement the two dimensional steady-state analysis completed recently by several authors (Szafer and Stone, 1989; Kirczenow, 1989 and references therein).

Hereafter, we limit our numerical simulations to the diffraction of a gaussian wavepacket through a narrow aperture in a potential wall of finite height. The thickness d and length w of the slit are assumed to be 100Å and 200Å respectively, which corresponds to an aspect ratio a = d/w equal to 0.5. The choice of such a narrow slit was imposed to make the problem tractable numerically without having to use an excessive number of grid points. However, the time evolution of the diffracted beam shows interesting features which could still be present in the electron diffraction through a more realistic split gate realizable with present-day technology (Thornton and coworkers, 1986; Zheng and coworkers, 1986). The use of electron diffraction through a split-gate was recently proposed by Kriman and coworkers (1988) for interesting device applications.

The Numerical Approach

Recently, Ancilatto and coworkers (1989) have developed a method to solve the multi-dimensional Schroedinger equation based on the Chebychev polynomial expansion of the time evolution operator. Each term in this expansion is calculated using FFT computations. The subsequent effort to calculate the quantum-mechanical wavefunction $\psi(t)$ scales roughly as $\sim MN^d \ln N$, where M is the number of terms in the Chebychev expansion, d is the number of space dimensions and N^d is the total number of grid points. One drawback of this approach is that it requires the use of uniformly spaced grid points. Another drawback of the Chebychev expansion of the time-evolution operator is that, being a truncated series expansion, it does not conserve the unitarity of the time-evolution operator unless the truncated series converges. As a consequence, energy and norms of the wavefunction need not be conserved at any time step.

In this paper, we used an alternative approach to solve the time-dependent Schroedinger equation based upon a finite-difference solution procedure to solve a set of coupled equations governing the real and imaginary components of the wavefunction. The use of Crank-Nicolson differencing scheme insures conservation of the norm of the wavefunction at all times. The resulting coupled difference equations are then solved using a block alternating direction implicit (ADI) technique following the scalar ADI development of Douglas and Gunn (1964). Recently, a similar algorithm has been used by Barker (1989) to study the wavepacket propagation through Aharanov-Bohm rings. The technique can be used with non-uniform grid spacings, allows for an explicit time-dependence of the potential energy profile and can readily be extended to include the presence of a spatially varying effective-mass, of an external magnetic field and to three dimensional configurations (Cahay and coworkers, 1989).

Numerical Examples

As an illustrative example, we consider the diffraction of a two-dimensional Gaussian wavepacket through a narrow constriction such as the one shown in the upper left frame of Fig. 1. The subsequent frames show the time-evolution of the wavepacket impinging on a 100 Å wide potential barrier of height 0.06eV containing a 200 Å wide slit where the potential is assumed to be zero. The slit is disposed symmetrically with respect to the center of the simulation domain, a square of dimensions 3,000Å x 3,000Å. Our numerical simulations were performed using a non-uniform grid spacing with 69 and 79 grid points in the x and y directions respectively. In an actual split-gate, the slit can actually be varied from a few hundred to a few thousand angstroms while sweeping the gate voltage. The potential in the 2D electron gas is also different from the sharp potential wall assumed in the present work. We will comment on this point in our conclusion section. The initial wavepacket is assumed to be

$$\psi(x,y,t=0) = \left[1/\sqrt{\pi\sigma^2}\right] \exp[ik_0 x] \exp[-\{(x-x_0)^2 + (y-y_0)^2\}/2\sigma^2] \tag{1}$$

Where (x_0, y_0) are the coordinates of the wavepacket center and σ is equal to 100Å. Finally, the electron wavevector k_0 is 0.0094 Å^{-1} (with this value of k_0, a free electron ($m^* = 0.067m_0$) travels a length of 1,000Å in 0.6 ps). The average kinetic energy of an electron in the state (1) is about 10 meV for the values of the parameters listed above and is therefore about 1/6th of the barrier height. A Fermi energy of about 10 meV corresponds to a typical 2D electron gas sheet density of ~ 3 x 10^{11} cm^{-2}. In Fig. 1, the frames are taken at various physical time steps equal to a multiple of 0.05 ps (which is equal to 10 times the computational time step). We have plotted the logarithmic contour plots ($\log_{10}$) of the probability density $|\psi(t)|^2$. For $k_0 = 0.0094$Å^{-1}, the electron De Broglie wavelength ($\lambda = 2\pi/k_0$) is about 660Å and therefore bigger than the slit width. Consequently, even though part of the wavepacket ($\geq 35\%$) is transmitted on the right-hand side of the wall at time t$\geq$0.25ps, no diffraction lobes are detectable in the transmitted waveform. For an electron with an incident energy of 45 meV (such an electron is far in the tail of the Fermi distribution in the previous example), $k_0 = 0.0265$Å^{-1}, which is close to the diffraction condition, $k_0 \approx 2\pi/w$. In that case, as can be seen in Fig. 2, there is actually a diffraction of the electron wavepacket through the slit as indicated by the existence of three distinct lobes in the contour plots of the charge density profiles in the immediate vicinity (on the right) of the slit at time t equal 0.15 ps. However, for later time steps, the diffraction lobes actually smear into a main one which (for this specific case) precludes the actual observation of a Fraunhoffer-like diffraction pattern far from the slit. In fact, the diffraction condition, $k_0 \approx 2\pi/w$, is only met by a small fraction of the various Fourier components of the wavepacket incident on the slit. As a result, the diffraction pattern cannot be as sharp as in the case of an incident plane wave, which is the idealistic situation equivalent to the one used in optics to study light diffraction through a narrow slit. More numerical simulations involving modification of the shape of the wavepacket, potential walls and slit dimensions and extension of the simulation domain need to be performed however before determining if an appropriate set of parameters can eventually lead to a Fraunhoffer-like diffraction pattern far from the slit. The overall shape of the wavepacket behind the slit changes drastically while varying the direction of incident wavevector k_0. This is illustrated in Fig. 3 where the wavepacket with a kinetic energy equal to 10 meV was assumed to be incident at a 45° angle on the slit. This example stresses the importance of collimating the electron beam in order to observe a diffraction pattern with different lobes beyond the slit. This point was already stressed by Kriman and coworkers (1988) in their steady state analysis of the QUADFET.

The width of the potential barrier below the split gate is of primary importance in determining the diffraction pattern behind the split-gate. As a result of the finite width, the electron wavepacket suffer multiple reflections between the potential walls defined by the slit. The resulting spread in transit times introduces additional structure in the diffracted electron beam. For instance, the time frame t = 0.35 ps in Fig. 2 above clearly shows the presence of successive maxima in the transmitted part of the wavepacket in the direction perpendicular to the potential wall. Finally, additional lobes can be clearly seen in the reflected part of the wavepacket, i.e., on the left side of the slit. Such "Back diffraction" would probably also exist while considering realistic split-gate and it could be eliminated by appropriately collimating the incident electron beam.

Conclusions

Preliminary results describing the time-evolution and diffraction of a wavepacket through a narrow slit in a finite potential wall have been reported. In order to save computational time, the size of the slit was chosen to be smaller than the actual split-gate realized with present day technology (Thornton and coworkers (1986); Van Wees and coworkers (1986). Several interesting features in the time evolution of the wavepacket have however been pointed out: (1) A Fresnel-like diffraction pattern is seen at early time once the diffraction requirement is

satisfied, i.e., $k_0 = 2\pi/w$ (k_0 being the wavepacket incident wavenumber and w being the slit width); (2) Due to the finite thickness of the slit (in our simulation, the aspect ratio of the slit is equal to 0.5), the main diffracted lobe is found to be fragmented in the direction perpendicular to the slit aperture as a result of the multiple reflections through the slit of finite length; (3) No Fraunhoffer-like diffraction pattern is seen far from the slit even when the slit width is comparable to the electron De Broglie wavelength.

One should emphasize that the Fresnel-like diffraction pattern observed in our simulations is an artifact of the use of infinitely sharp corners in the potential energy profile. A more realistic simulation would require self-consistent calculations of the potential energy profile below a split-gate such as reported by Laux and coworkers (1988). Even with appropriate collimation of a nearly monochromatic incident electron beam, the rounding of the corners below the split-gate on the scale of the wavelength of the incident electron could be the dominant source of suppression of the diffraction pattern at low temperatures.
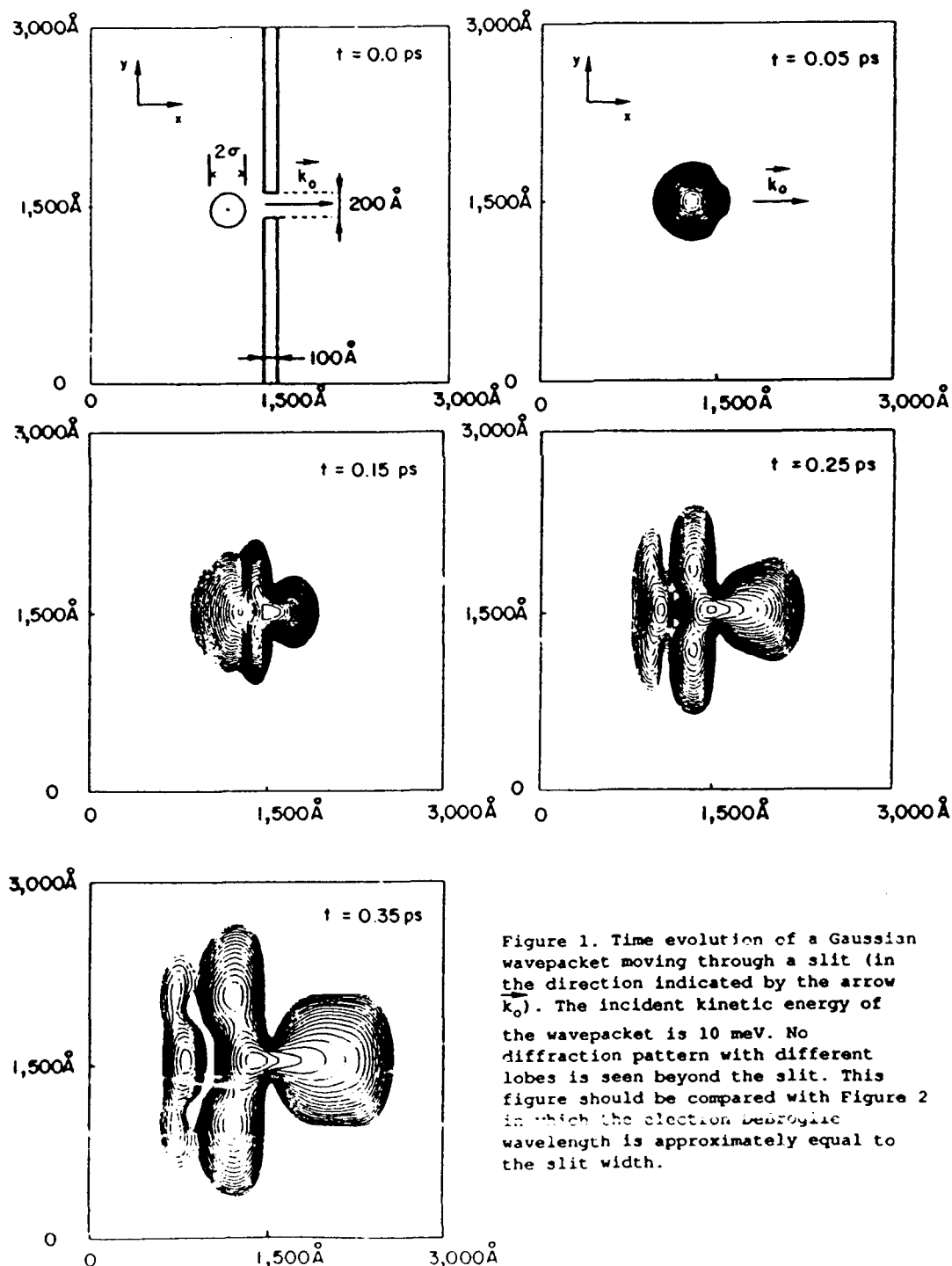

Figure 1. Time evolution of a Gaussian wavepacket moving through a slit (in the direction indicated by the arrow $\overrightarrow{k_0}$). The incident kinetic energy of the wavepacket is 10 meV. No diffraction pattern with different lobes is seen beyond the slit. This figure should be compared with Figure 2 in which the electron DeBroglie wavelength is approximately equal to the slit width.
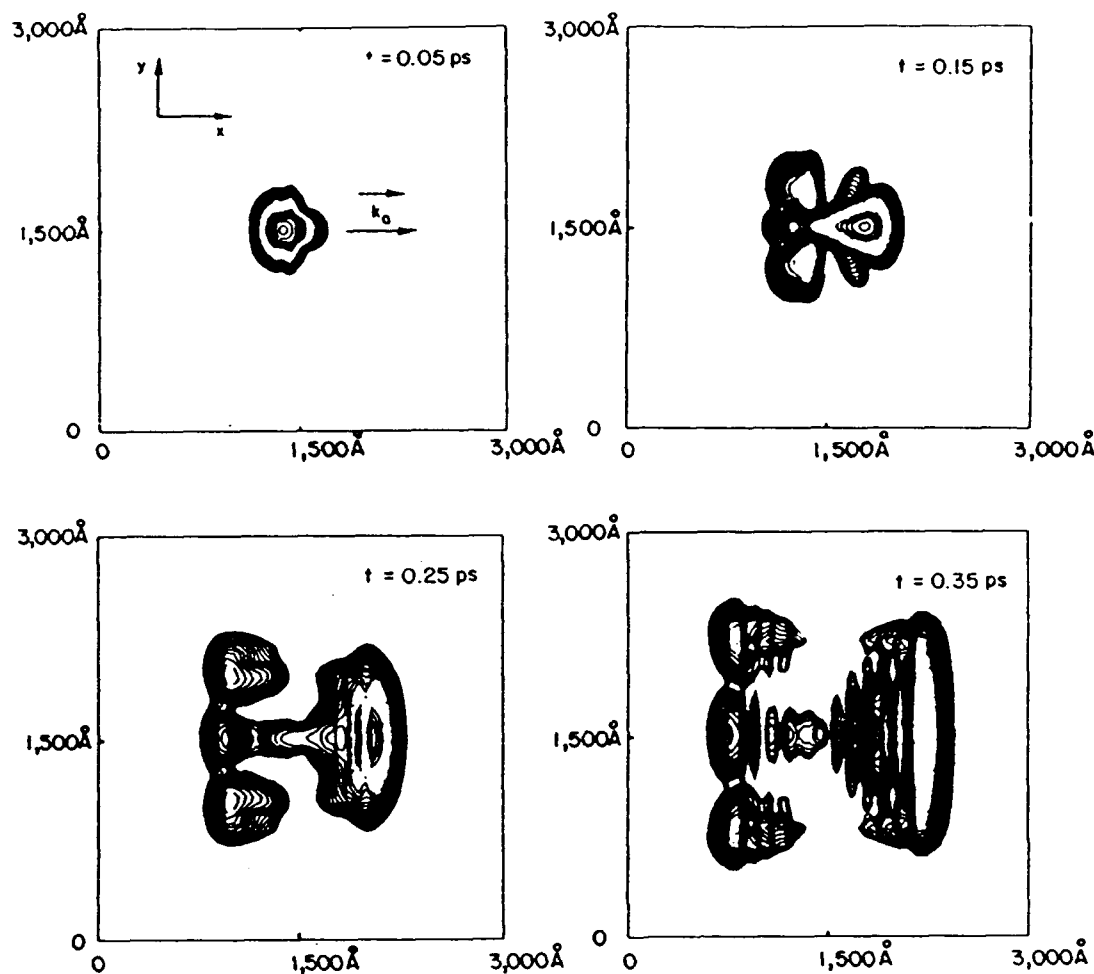
Figure 2.   Same as Figure 1 for a wavepacket with an incident energy equal
to 45 meV. Note the formation of diffraction lobes after crossing the slit
which can clearly be seen at time t = 0.15ps. At time t = 0.35ps, the main
transmitted lobe of the diffracted wavepacket is clearly segmented into
various fragments. These are due to multiple reflections of the wavepacket
inside the slit which has finite length. As a result, different parts of the
wavepacket are delayed further in time leading to successive transmitted lobes.
Spurious effect due to reflection at the boundaries of the simulation domain
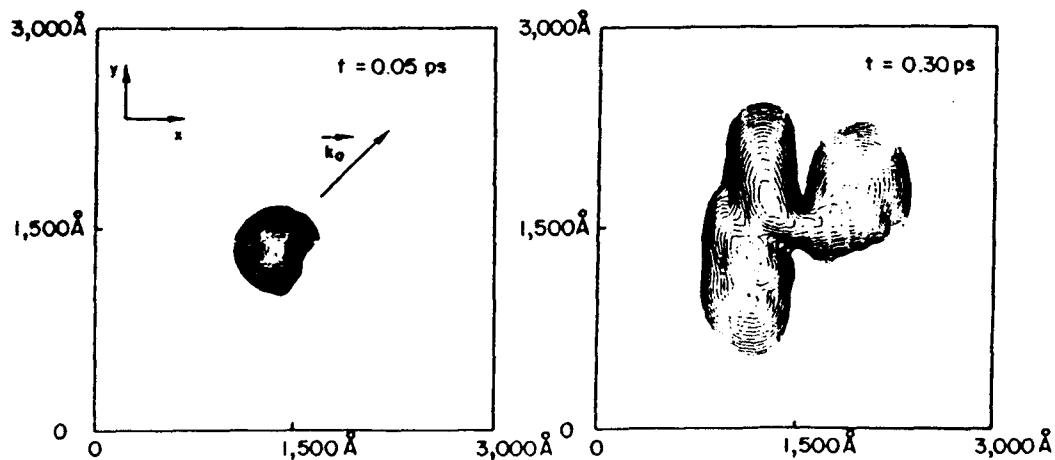are only seen for t = 0.5ps.

Figure 3.   Same as Figure 1 while shooting at an angle of 45°
on the slit. The incident kinetic energy is 10 meV.

References

Ancilatto, F., Selloni, A., Xu, L. F., and E. Tosatti, (1989). Phys. Rev. B. 39, 2993.

Bandyopadhyay, S., Bernstein G. H., and W. Porod (1989). Proc. First Conference on Nanostructure Physics and Fabrication, Ed. by M. A. Reed and W. P. Kirk, Academic Press, Boston.

Barker, J. R. (1989). Proc. First Conference on Nanostructure Physics and Fabrication, Ed. by M. A. Reed and W. P. Kirk, Academic Press, Boston; Also, Proc. of San Miniato Workshop on Quantum Transport (1987), San Miniato, Italy.

Cahay, M., Kreskovsky, J. P., and H. L. Grubin (1989). Unpublished.

Douglas, J., and J. E. Gunn (1964). Numer. Math, 6.

Goldberg, A., Shey, H. M. and J. C. Swartz, (1967). Am. J. Phys. 35, 177.

Hauge, E. M. and J. A. Stovneng, (1989). Preprint.

Ihayama, Y., Saku, T. and Y. Horikoshi (1989).Phys. Rev. B. 39, 5535.

Kirczenow, G., (1989). Phys. Rev. Lett. 62, 2993 and references therein.

Kriman, A. M., Bernstein, G. H., Haukness, B. S., and D. K. Ferry (1988). Proc. 4th Intl. Conf. on Superlattices. Microstructures and Microelectronics, Trieste, Italy.

Laux, S. E., Frank, D. J., and F. Stern (1988). Surface Science. 196, 101.

Szafer, A. and A. D. Stone (1989). Phys. Rev. Letters. 62, 300.

Thornton, T. J., Pepper, M., Ahmed, H., Andrews, D., and G. J. Davies (1986) Phys. Rev. Letters 56, 1198.

Van Wees, B. J., Van Houten, H., Beenakker, C. W. J., Williamson, J. G., Kouwenhoven, L. P., Van der Marel, D., and C. T. Foxon (1988). Phys. Rev. Letters. 60, 848.

Zheng, H. Z., Wei, H. P., Tsui, D. C., and G. Weimann (1986). Phys. Rev. B. 34, 5635.

Acknowledgement

Numerical Solution of the Equation of Motion of the Coordinate Representation Density Matrix

T. R. Govindan, H. L. Grubin and F. J. deJong
Scientific Research Associates, Inc.
Glastonbury, Connecticut 06033-1058

ABSTRACT

Mathematically well posed problems have been formulated and numerical schemes have been implemented for solving the equation of motion of the density matrix. Equivalent first order systems of equations have been derived for the equation of motion of the density matrix and these systems are solved rather than the second order equation of motion. The advantage of the approach is that suitable forms of boundary conditions can be inferred from the characteristic directions of the first order system. Further, the first order system can be solved rapidly by a 'method of characteristics' based numerical scheme that has been developed. Typical solutions, which are discussed, require about 20 seconds of Cray-XMP CPU time and demonstrate the accuracy and efficiency of the solution procedure.

INTRODUCTION

Frensley [1] was the first to suggest that solutions to the equation of motion of the density matrix in the coordinate representation could be an important tool in understanding the physics of quantum well and quantum barrier devices. In [2] we described typical results obtained from a solution algorithm we have developed for the equation of motion of the density matrix. In this paper we describe essential features of the solution algorithm.

SOLUTION PROCEDURE

The equation of motion of the density matrix is [1]:

$$\frac{\partial}{\partial t} \rho(x,x',t) = \frac{i\hbar}{2m}\left[\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial x'^2}\right]\rho(x,x',t)$$
$$- \frac{i}{\hbar}[v(x,t)-v(x',t)]\rho(x,x',t) \quad (1)$$

where ρ is the density matrix, m is the effective mass, and v is the potential energy. We focus attention here on obtaining time-independent solutions of equation (1).

For convenience of solution and determining suitable forms of boundary conditions, equation (1) is written in the form of a coupled first order system of equations:

$$u(x,x') + \frac{i\hbar}{2m}\left[\frac{\partial\rho}{\partial x} + \frac{\partial\rho}{\partial x'}\right] = 0 \quad (2)$$

$$\frac{\partial\rho}{\partial t} + \left[\frac{\partial u}{\partial x} - \frac{\partial u}{\partial x'}\right] + \frac{i}{\hbar}[v(x)-v(x')]\rho=0 \quad (3)$$

Equation (2) is the definition of $u(x,x')$ and equation (3) is equation (1) written in terms of u and ρ. The characteristic directions for the system of equations (2) and (3) are

$$\eta = (x + x')/2 = \text{constant} \quad (4)$$

$$\zeta = (x - x')/2 = \text{constant} \quad (5)$$

In terms of the characteristic directions η and ζ, equations (2) and (3) can be written as

$$u(x,x') + \frac{i\hbar}{2m}\frac{\partial\rho}{\partial\eta} = 0 \quad (6)$$

$$\frac{\partial\rho}{\partial t} + \frac{\partial u}{\partial\zeta} + \frac{i}{\hbar}[v(x)-v(x')]\rho=0 \quad (7)$$

Suitable boundary conditions for equations (6) and (7) are the specification of ρ and u along the boundary x'=0 and the specification of u along the boundary x=L, where L is the length of the device. Along the boundary x=0, ρ is specified as the complex conjugate of $\rho(x,0)$, since ρ is hermitian, and u is computed from the outgoing characteristic equation (7). Along the boundary x=L, ρ is computed from the outgoing characteristic equation (6).

An alternate system of first order equations can be formulated in terms of the current matrix.

$$j(x,x') + \frac{i\hbar}{2m}\frac{\partial\rho}{\partial\zeta} = 0 \quad (8)$$

$$\frac{\partial\rho}{\partial t} + \frac{\partial j}{\partial\eta} + \frac{i}{\hbar}[v(x)-v(x')]\rho=0 \quad (9)$$

Equations (8) and (9) have the same characteristic directions η and ζ as equations (2) and (3). Suitable boundary conditions for equations (8) and (9) are the specification of ρ and j along the boundary x'=0 and the specification of ρ along the boundary x=L. Along the boundary x=0, j is specified as the complex conjugate of $j(x,0)$ and ρ is computed from the outgoing characteristic equation (8). Along the boundary x=L, j is computed from the outgoing characteristic equation (9). Both sets of the first order system of equations, equations (6)-(7) and equations (8)-(9), are useful in applications since they allow different forms of boundary conditions. Both sets of equations can be solved by the same numerical procedure.

The overall solution procedure consists of solving the first order system of equations as an initial boundary-value problem starting from conditions along the line x'=0 and marching to the line x'=L using the method of characteristics. A characteristic net for the equation of motion of the density matrix can be constructed a-priori from grid points of a uniform square grid. A discrete form of equations (6) and (7) can be written on this grid as [Figure 1].

$$[u]_{av} + \frac{i\hbar}{2m} \cdot \frac{\rho(i,j) - \rho(i-1,j-1)}{\Delta\eta} = 0 \quad (10)$$

$$\left[\frac{\partial\rho}{\partial t}\right]_{av} + \frac{u(i+1,j-1) - u(i,j)}{\Delta\zeta}$$
$$\quad (11)$$
$$+ \frac{i}{\hbar} [v(x) - v(x')]_{av}[\rho]_{av} = 0$$

where $[\cdot]_{av}$ represents an average over the grid cell. Depending upon the form of averaging chosen, equations (10) and (11) form a system of 2x2 block tridiagonal or block diagonal algebraic equations that can be easily solved at x'=j from known values at x'=j-1. Thus, the solution procedure can be marched from boundary conditions at x'=0, in steps along x', to x'=L. A similar procedure can be utilized for equations (8) and (9). Self-consistency is included in the analysis by iterating the solution of the density matrix equation with the solution of Poisson's equation to convergence.

Figure 1. Discretization of the Equations

RESULTS

The equilibrium distribution of potential and density for a double barrier structure of [3] were computed. Computations were carried out on a uniform grid of 300 x 300 points. Maxwellian distribution of density was assumed at the boundary. Figure 2 shows the self-consistent potential distribution in the structure. Figure 2 also shows the density distribution with local accumulation of charge in the quantum well due to tunneling. Figure 3 is a plot of the density matrix on the (x,x') plane, showing Maxwellian distribution near the boundaries, depletion of charge in the vicinity of the barriers, and accumulation in the quantum well. For brevity, only one typical result has been presented here. The solution procedure has been used to compute distributions in a variety of structures.

CONCLUSIONS

The ability to compute solutions of the equation of motion of the density matrix provides a computational tool to examine fundamental issues associated with mesoscopic structures. Of particular relevance is that the time independent computations took about 20 CPU seconds on a Cray X-MP and thereby provide a new efficient means of exploring the physics of quantum structures.

Figure 2. Potential and Density vs. Distance

REFERENCES

[1] W. R. Frensley, J. Vac. Sci. Technol, B3, 1261 (1985).

[2] T. R. Govindan, H. L. Grubin, and F. J. deJong, Proceedings of the Workshop on Computational Electronics, Beckmann Institute, 21-22 May 1990.

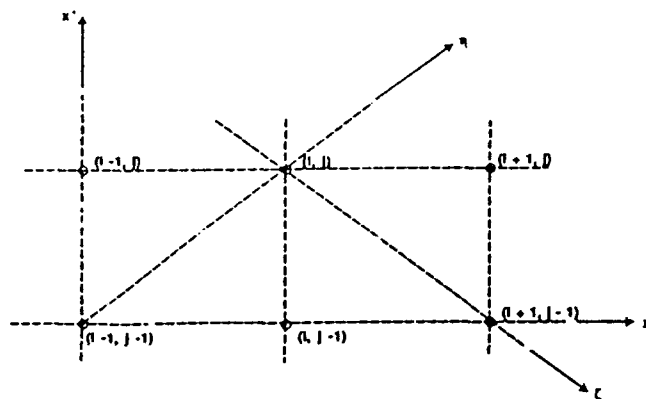[3] N. C. Kluksdahl, A. M. Kriman, D. K. Ferry, and C. Ringhofer, Phys. Rev. B39, 7720 (1989).
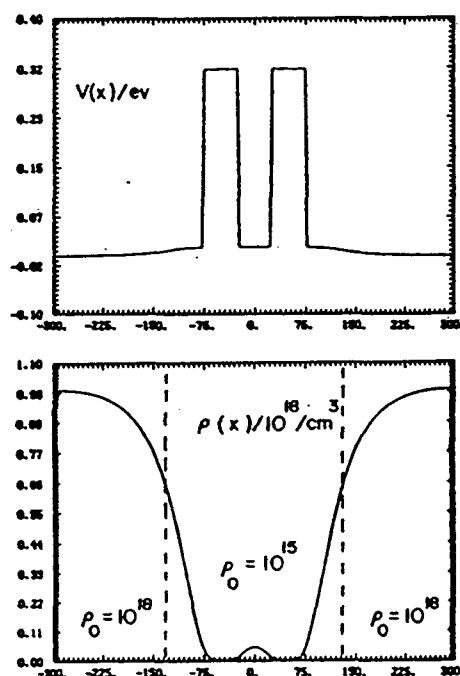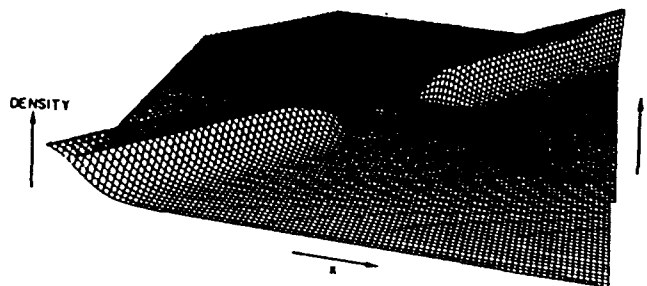
Figure 3. Density Matrix on the (x,x') Plane

SPACE-CHARGE EFFECTS IN COMPOSITIONAL AND EFFECTIVE-MASS SUPERLATTICES

M. Cahay[1], M. A. Osman[1], H. L. Grubin[1]

Scientific Research Associates, Inc.
Glastonbury, CT   06033

M. McLennan[2]

School of Electrical Engineering
Purdue University
West Lafayette, Indiana   47907

I.   INTRODUCTION

In the Effective-Mass Superlattice (EMSL) proposed by Sasaki [1,2], the effective-mass of electrons (or holes) is changed periodically and the conduction (for electrons) or valence band (for holes) is supposedly aligned. This eliminates the potential discontinuities between the respective superlattice layers which are present in doping and composiｃional superlattices. Previous work [3] neglecting space-charge effects has shown the threshold voltage for negative differential resistance (NDR) in EMSL to be much lower ($\approx$10mV) and the current density

$(10^4-10^5 \text{ A/cm}^2)$ much higher than in resonant tunneling diodes (RTD). Hereafter, we report preliminary self-consistent calculations of space-charge effects in EMSL and compositional superlattices under zero bias condition.

II. THEORY

We use the technique described in ref. [4] to calculate quantum-mechanically the electron charge density profile in various types of superlattices (see fig. 1). Under the assumption of ballistic transport, the Schroedinger equation throughout the entire structure is solved using the scattering-matrix approach [5] for each electron impinging from the contacts, with the usual boundary conditions for plane-wave solutions, requiring everywhere continuity of the wave function and its first derivative divided by the electron effective mass. Once the Schroedinger equation is solved for electrons impinging from the contacts, the electron density is then calculated for the two streams of electrons incident from the left and right contacts. The calculation for electron density and electrostatic potential are then performed iteratively, until the electrostatic potential converges to a final solution.
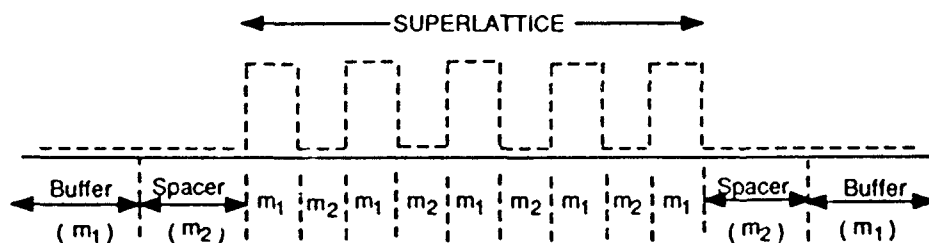

Fig. 1. Spatial variation of the conduction band edges of the device studied in the text. The dashed (full) curve describes a compositional (EMSL) superlattice respectively. Only the buffer regions are assumed to be doped (10^{18}cm^{-3}).

III. NUMERICAL EXAMPLES

The technique outlined above was applied to calculate self-consistently the electron density and conduction band energy profile of two types of superlattices whose configuration are shown in figure 1. The first superlattice consists of five regions of $In_{0.72}Ga_{0.28}As_{0.86}P_{0.14}$ with width 41Å and effective mass $m_1 = 0.039m_0$ separated by InP regions of width 29.34Å and effective mass $m_2 = 0.073m_0$. The current-voltage characteristic of this EMSL was calculated in ref. [3] neglecting space-charge effects. As shown in ref. [1], this special choice of layer thicknesses leads to equal energy gaps between the EMSL energy subbands.

This EMSL is sandwiched between two undoped InP spacer layers (50Å) contacting two highly doped ($10^{18}cm^{-3}$) InP buffer regions (500Å). The second superlattice is a compositional superlattice obtained by replacing the $In_{0.72}Ga_{0.28}As_{0.86}P_{0.14}/InP$ regions in the EMSL described above by $Al_{0.3}Ga_{0.7}As/GaAs$ respectively ($m_1 = 0.0919m_0$ and $m_2 = 0.067m_0$). The conduction band discontinuity between $Al_{0.3}Ga_{0.7}As/GaAs$ is taken to be 0.209 eV (see fig. 1).

Figure 2a) shows the self-consistent charge density profile in the two different superlattices under zero bias condition. Also shown for comparison is the charge density profile obtained without iteration, i.e., using the conduction band energy profile such as shown in figure 1. As a result of the doping gradients, internal contact potentials are important (while different) in both types of superlattices as can be seen in figure 2b). This leads to an upward shift of the conduction band energy profile in both superlattices resulting in a substantial decrease of the charge density, when calculated self-consistently, from their non self-consistent values. As shown in fig. 2a), for identical doping concentration in the buffer regions, the charge density is about one order of magnitude higher in the EMSL than in the compositional superlattice.

a)

T = 300 K

SELF-CONSISTENT
--- NOT SELF-CONSISTENT

EMSL

n'-n-n'

COMPOSITIONAL

ELECTRON CHARGE DENSITY (cm^{-3})

POSITION (cm)(x 10^{-4})

b)

T = 300 K

COMPOSITIONAL

n'-n-n'
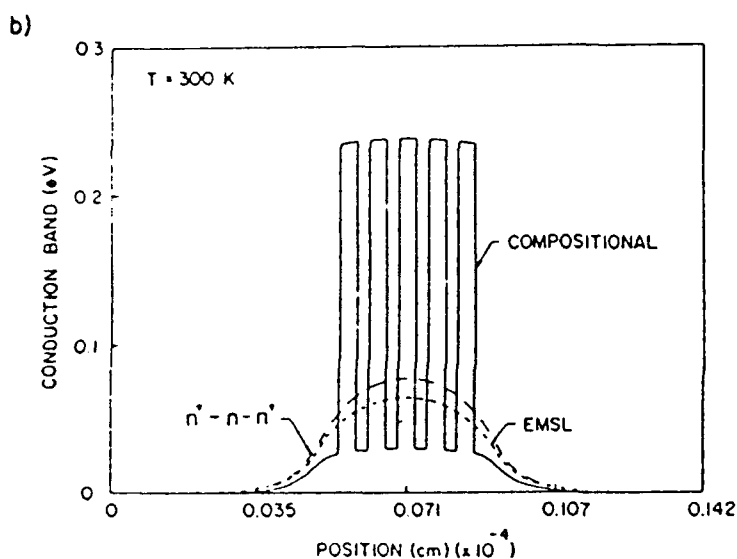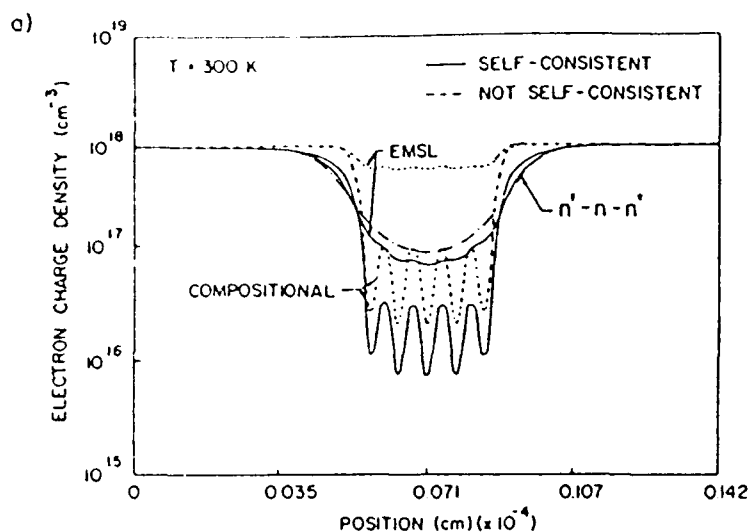
EMSL

CONDUCTION BAND (eV)

POSITION (cm)(x 10^{-4})

Fig. 2. Self-consistent and non self-consistent
(a) charge density profiles, (b) conduction band
energy profile in an EMSL and compositional
superlattice (see text and fig. 1) with identical
doping concentration in the buffer regions. The
dashed-dotted curves show the self-consistent
results for a structure in which the compositional
superlattice has been replaced by undoped GaAs
(n^+-n-n^+).

498

IV.   CONCLUSIONS

Our preliminary (zero bias) self-consistent calculations have shown that, in superlattices of similar dimensions, space-charge effects are more important in EMSL than in compositional super-lattices. The resulting increased overall capacitance of EMSL should therefore have a detrimental effect on their potentiality for high-speed device applications [3]. Indeed, inclusion of space-charge effects in RTD has been shown to reduce (up to 50%) their calculated peak current and peak-to-valley ratio and to shift to a higher voltage their NDR [4] (to an extent increasing with the importance of space charge). A careful investigation of the superiority of EMSL over RTD for achieving low-power and ultrafast bistable switches must therefore await a detailed self-consistent calculation of their current-voltage characteristic.

REFERENCES

1. A. Sasaki, Phys. Rev. B 30, 7016 (1984).
2. A. Sasaki, Surf. Sci. 174, 624 (1986).
3. A. Aishima and Y. Fukushima, J. Appl. Phys. 61 (1), 249 (1988); Electronic Letters 24, 65 (1988).
4. M. Cahay, M. McLennan, S. Datta and M. S. Lundstrom, Appl. Phys. Lett. 51 (10), 612 (1987).
5. M. Cahay, M. McLennan and S. Datta, Phys. Rev. B37, 10125 (1988).

DENSITY MATRIX COORDINATE REPRESENTATION NUMERICAL STUDIES OF QUANTUM WELL AND BARRIER DEVICES*

T. R. Govindan, H. L. Grubin and F. J. de Jong
Scientific Research Associates, Inc.
Glastonbury, Connecticut 06033-6058

INTRODUCTION

Frensley [1] was the first to suggest that solutions to the equation of motion of the density matrix in the coordinate representation, $\rho(x,x't)$ as an important tool in understanding of the physics of quantum well and quantum barrier devices. A mathematically well posed problem has been implemented here for solving this equation.

In one dimension, the equation of motion of the density matrix is:

$$(1) \quad \partial\rho(x,x',t)/\partial t = (i\hbar/2m)[\partial^2/\partial x^2 - \partial^2/\partial x'^2]\rho(x,x',t)$$
$$-(1/\hbar)[V(x,t)-V(x',t)]\rho(x,x',t)$$

We have solved the above equation using two separate solution procedures. The first procedure, for which only time independent results are discussed, involves the method of characteristics in which we obtain equivalent first order systems. For example, in one case, the coupled first order system is obtained after defining the current matrix:

$$(2) \quad j(x,x') + (i\hbar/2m)[\partial\rho/\partial x - \partial\rho/\partial x'] = 0$$

which transforms the density matrix equation into:

$$(3) \quad \partial\rho/\partial t + \partial j/\partial x + \partial j/\partial x' + (1/\hbar)[V(x)-V(x')]\rho = 0$$

Equations (2) and (3) represent an equivalent first order coupled system for obtaining the real and imaginary parts of the density and current matrices. With this formulation, boundary conditions require the specification of the density and the current along the boundaries $x' = 0$, $0 < x < L$, where L is the length of the device, the specification of current along $x = 0$, $0 < x' < L$, and the specification of density along $x = L$, $0 < x' < L$. Self-consistency is included in the analysis.

RESULTS

Equilibrium zero bias calculations: The first set of calculations represent equilibrium self-consistent calculations with Maxwellian boundaries. In equilibrium at the boundaries $\rho(x,x') = A \exp\{-[(x-x')/2\lambda]^2\}$ where $\lambda^2 = \hbar^2/2mk_bT$. Figure 1 displays the diagonal elements of density and potential for an $N^+N^-N^+$ structure with the indicated densities. The point to note about this structure is that the barrier height cannot be obtained from the classical expression, $\exp-V(x)/k_bT$, a result that is consistent with [2] and forms part of the basis for the introduction of the quantum potential in [3]. It is noted that as the length of the structure as well as the undoped region are increased, the potential increases and is accompanied by a decrease in carrier density. The result when a single barrier is centrally placed within the structure is shown in Figure 2. For this calculation the barrier height is approximately 10mev less than that of Figure 1, and suggests that simple estimates of the barrier in mesoscopic structures are likely to be inaccurate. Figures 3 and 4 display the equilibrium calculations for a double barrier structure of [4] who obtained solutions using the Wigner formulation. The agreement for this zero bias calculation is remarkable. Note that for the structure with doping up to the spacer layers, there is a region of local charge accumulation. The energy of the carriers is near and about the resonant level.
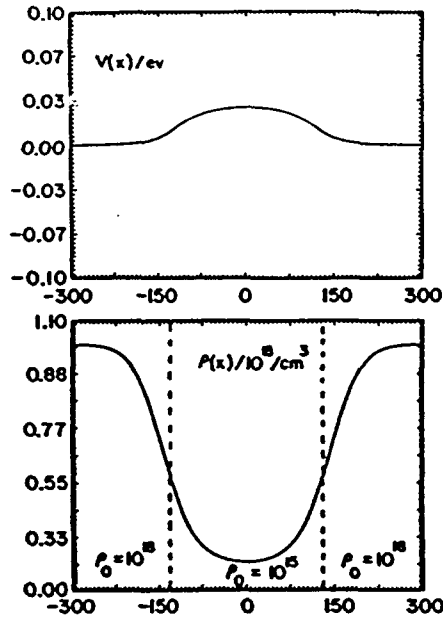
216

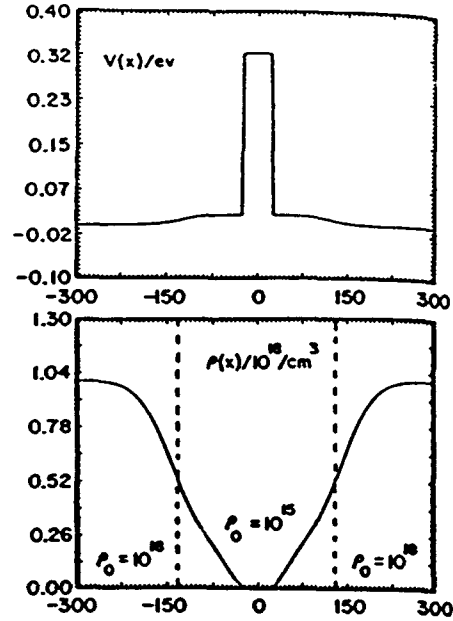Fig. 1. Potential and Density vs. Distance.

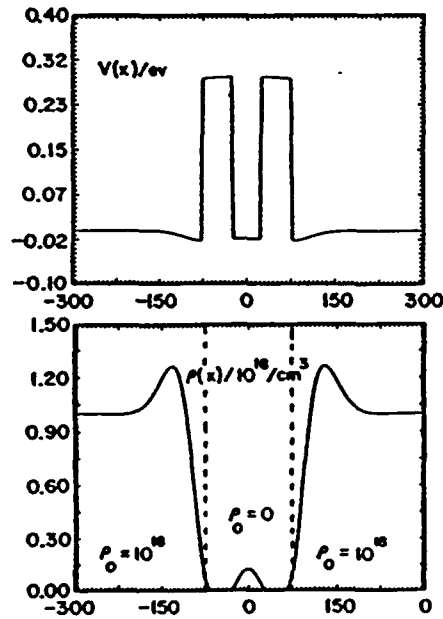
Fig. 2. Potential and Density vs. Distance.
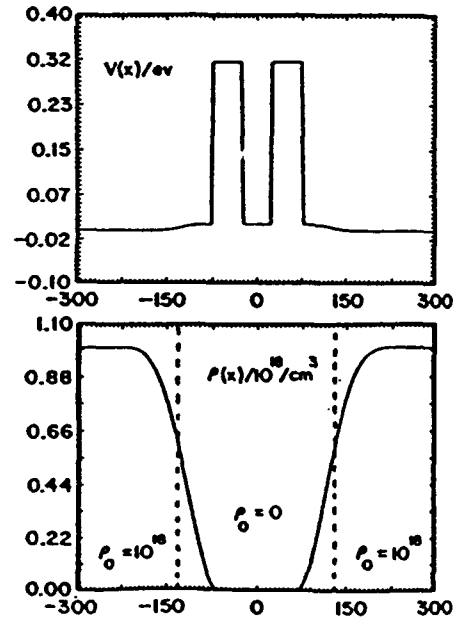

Fig. 3. Potential and Density vs. Distance.
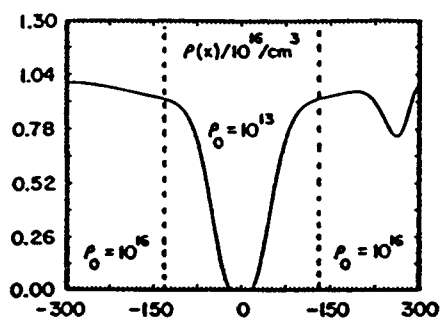

Fig. 4. Potential and Density vs. Distance.
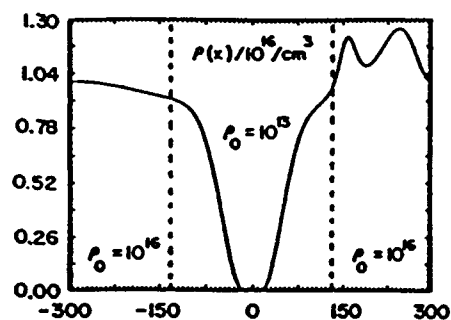
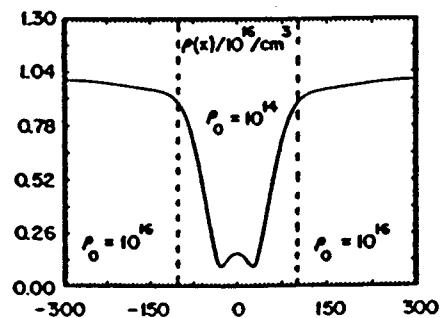Fig. 5. Density vs. Distance.
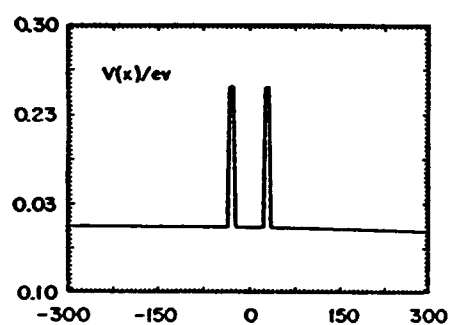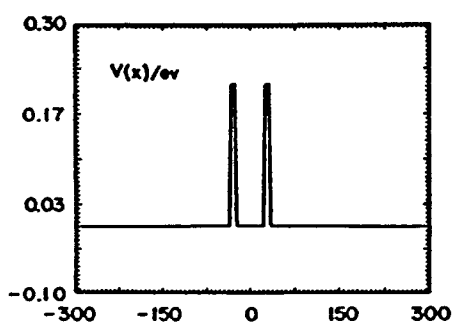
Fig. 6. Density vs. Distance.
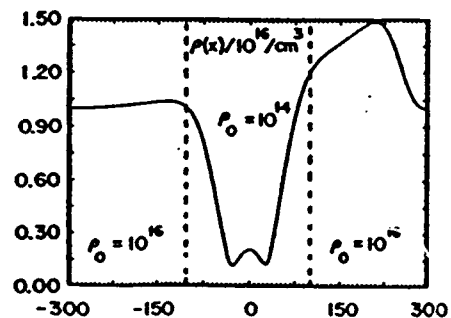
Fig. 7. Potential and Density vs. Distance.

Fig. 8. Potential and Density vs. Distance.

218

The calculations of Figures 1 through 4, assume Maxwellian boundary conditions on both the up-and downstream regions. If the distribution of quantum states is different for the two boundaries, it is expected that the corresponding distribution of carriers will be difference. Several calculations of this type were performed under the assumption of Maxwellian distributions at different temperatures. No physical significance is attributed to this temperature difference; rather the calculations are displayed to shown the significance of the boundary conditions. Figure 5 shows the result in which the anode boundary temperature is higher than that of the upstream contact. Figure 6 corresponds to the case when the temperature is lower. Note that in one case there is local accumulation at the downstream region. In the second case there is local depletion.

Finite bias calculations: In a companion paper, [5], the moments of the Wigner function are used to examine transport in a double barrier heterostructure. The barrier height is 21mev, and the barrier widths are a narrow 5 Å. The bias across the device [5] in the study is 8mev. The same structure in the structure was also examined using the density matrix. The equilibrium distribution of carrier is displayed in Figure 7 and the nonequilibrium distribution is displayed in Figure 8. Note that in Figure 8, current injection with an entering velocity of 4×10^6 cm/s was assumed. The results are in qualitative agreement with that of [5] particularly with respect to the almost linear variation of potential across the structure, although we note the high carrier density at the downstream boundary. As indicated earlier this value is determined by the spectra of energies at the downstream contact.

Time dependent calculations: A non self-consistent time dependent problem was solved using a different algorithm. The structure was a 650Å long GaAs element with two 50Å thick barriers separated by a 50Å region. At time t = 0, a uniform density of carriers was created. While this initial distribution is not physical, we point out that within the first 10fs of the calculation, the carriers are removed from the barriers and settle within the well and outside of the barriers. Subsequent time development shows that as the carriers outside of the double barrier move toward the contacts there is a reduction of charge within the well. Upon reflection at the contacts the charge in the well increases.

CONCLUSIONS

The above calculation demonstrates the ability of the density matrix to examine fundamental issues associated with mesoscopic structures. Of particular relevance is that the time independent calculations took less than 10 CPU seconds and thereby provide a new efficient means of exploring the physics of quantum structures.

REFERENCES

[1]   W. R. Frensley, J. Vac. Sci. Technol, B3, 1261 (1985).

[2]   E. P. Wigner, Phys, Rev. 40, 749 (1932).

[3]   H. L. Grubin and J. P. Kreskovsky, Solid State Electronics, 32, 1071, (1989).

[4]   N. C. Kluksdahl, A. M. Kriman, D. K. Ferry and C. Ringhofer, Phys. Rev. B 39, 7720 (1989)

[5]   D. L. Wollard, M. A. Stroscio, M. A. Littlejohn, R. J. Trew, and H. L. Grubin, Proceedings of the Workshop on Computational Electronics, Beckmann Institute 21-22 May 1990.

DENSITY MATRIX COORDINATE REPRESENTATION NUMERICAL STUDIES OF QUANTUM WELL AND BARRIER DEVICES[*]

T. R. Govindan, H. L. Grubin and F. J. de Jong
Scientific Research Associates, Inc.
Glastonbury, Connecticut 06033-6058

INTRODUCTION

Frensley [1] was the first to suggest that solutions to the equation of motion of the density matrix in the coordinate representation, $\rho(x,x't)$ as an important tool in understanding of the physics of quantum well and quantum barrier devices. A mathematically well posed problem has been implemented here for solving this equation.

In one dimension, the equation of motion of the density matrix is:

(1) $\partial\rho(x,x',t)/\partial t = -(i\hbar/2m)[\partial^2/\partial x^2 - \partial^2/\partial x'^2]\rho(x,x',t) - (1/\hbar)[V(x,t)-V(x',t)]\rho(x,x',t)$

We have solved the above equation using two separate solutions procedures. The first procedure, for which only time independent results are discussed, involves the method of characteristics in which we obtain equivalent first order systems. For example, in one case, the coupled first order system is obtained after defining the current matrix

(2) $\cdot j(x,x') + (i\hbar/2m)[\ \partial\rho/\partial x - \partial\rho/\partial x'] = 0$

which transforms the density matrix equation into:

(3) $\partial\rho/\partial t + \partial j/\partial x + \partial j/\partial x' + (1/\hbar)[V(x)-V(x')]\rho = 0$

Equations (2) and (3) represent an equivalent first order coupled system for obtaining the real and imaginary parts of the density and current matrices. With this formulation, boundary conditions require the specification of the density and the current along the boundaries $x' = 0$, $0 < x < L$, where L is the length of the device, the specification of current along $x = L$, $0 < x' < L$, and the specification of density along $x = L$, $0 < x' < L$. Self-consistency is included in the analysis.

RESULTS

Equilibrium zero bias calculations: The first set of calculations represent equilibrium self-consistent calculations with Maxwellian boundaries. In equilibrium at the boundaries $\rho(x,x') = A \exp\{-[(x-x')^2/2\lambda_t]^2$ where $\lambda_t^2 = \hbar^2/2mk_bT$. Figure 1 displays the diagonal elements of density and potential for an $N^+N^-N^+$ structure with the indicated densities. The point to note about this structure is that the barrier height cannot be obtained from the classical expression, $\exp-V(x)/k_bT$, a result that is consistent with [2] and forms part of the basis for the introduction of the quantum potential in [3]. It is noted that as the length of the structure as well as the undoped region are increased, the potential increases and is accompanied by a decrease in carrier density. The result when a single barrier is centrally placed within the structure is shown in Figure 2. For this calculation the barrier height is approximately 10mev less than that of Figure 1, and suggests that simple estimates of the barrier in mesoscopic structures are likely to be inaccurate. Figures 3 and 4 display the equilibrium calculations for a double barrier structure of [4] who obtained solutions using the Wigner formulation. The agreement for this zero bias calculation is remarkable. Note that for the structure with doping up to the spacer layers, there is a region of

local charge accumulation. The energy of the carriers is near and about the resonant level.

The calculations of Figures 1 through 4, assume Maxwellian boundary conditions on both the upper and downstream regions. If the distribution of quantum states is different for the two boundaries, it is expected that the corresponding distribution of carriers will be difference. Several calculations of this type were performed under the assumption of Maxwellian distributions at different temperatures. No physical significance is attributed to this temperature difference; rather the calculations are displayed to shown the significance of the boundary conditions. Figure 5 shows the result in which the anode boundary temperature is higher than that of the upstream contact. Figure 6 corresponds to the case when the temperature is lower. Note that in one case there is local accumulation at the downstream region. In the second case there is local depletion.

Finite bias calculations: In a companion paper, [5], the moments of the Wigner function are used to examine transport in a double barrier heterostructure. The barrier height is 21mev, and the barrier width's are a narrow 5 Å. The bias across the device [5] in the study is 8mev. The same structure in the structure was also examined using the density matrix. The equilibrium distribution of carrier is displayed in Figure 7 and the nonequilibrium distribution is displayed in Figure 8. Note that in Figure 8, current injection with an entering velocity of 4x10⁶ cm/s was assumed. The results are in qualitative agreement with that of [5] particularly with respect to the almost linear variation of potential across the structure, although we note the high carrier density at the downstream boundary. As indicated earlier this value is determined by the spectra of energies at the downstream contact.

Time dependent calculations: A non self-consistent time dependent problem was solved using a different algorithm. The structure was a 650Å long GaAs element with two 50Å thick barriers separated by a 50Å region. At time t = 0, a uniform density of carriers was created. While this initial distribution is not physical, we point out that within the well and outside of the barriers, carriers are removed from the barriers and settle within the well and outside of the barriers. Subsequent time development shows that as the carriers outside of the double barrier move toward the contacts there is a reduction of charge within the well. Upon reflection at the contacts the charge in the well increases.

CONCLUSIONS

The above calculation demonstrates the ability of the density matrix to examine fundamental issues associated with mesoscopic structures. Of particular relevance is that the time independent calculations took less than 10 CPU seconds and thereby provide a new efficient means of exploring the physics of quantum structures.

REFERENCES

[1] W. R. Frensley, J. Vac. Sci. Technol, B3, 1261 (1985).

[2] E. P. Wigner, Phys. Rev. 40, 749 (1932).

[3] H. L. Grubin and J. P. Kreskovsky, Solid State Electronics, 32, 1071, (1989).

[4] N. C. Kluksdahl, A. M. Kriman, D. K. Ferry and C. Ringhofer, Phys. Rev. B 39, 7720 (1989).

[5] D. L. Wollard, M. A. Stroscio, M. A. Littlejohn, R. J. Trew, and H. L. Grubin, Proceedings of the Workshop on Computational Electronics, Beckmann Institute 21-22 May 1990.
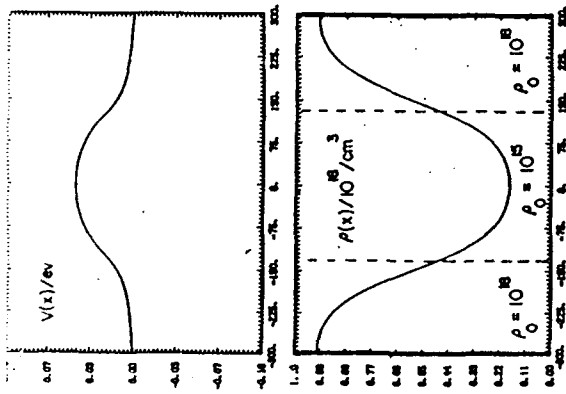
Fig. 6. Density vs. Distance.

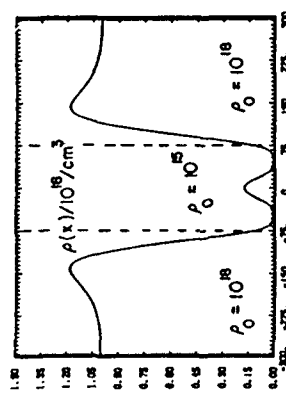Fig. 5. Density vs. Distance.

Fig. 8. Potential and Density vs. Distance.

Fig. 7. Potential and Density vs. Distance.
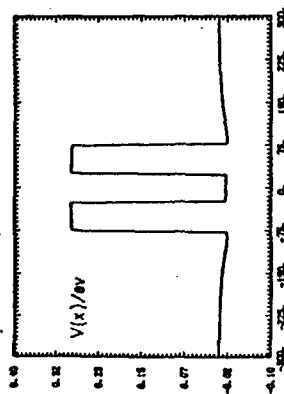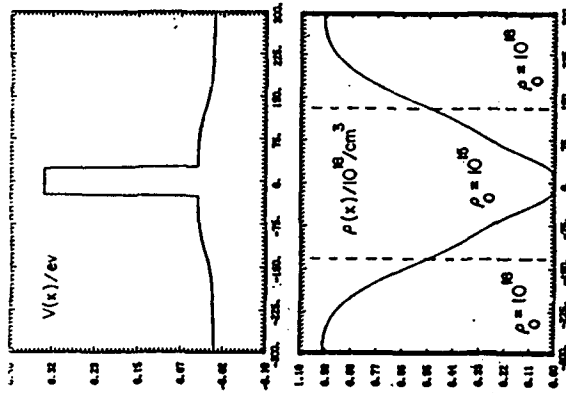
Fig. 2. Potential and Density vs. Distance.

Fig. 1. Potential and Density vs. Distance.
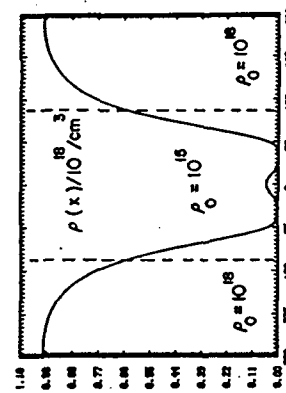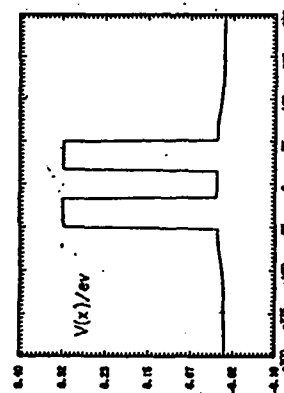
Fig. 4. Potential and Density vs. Distance.

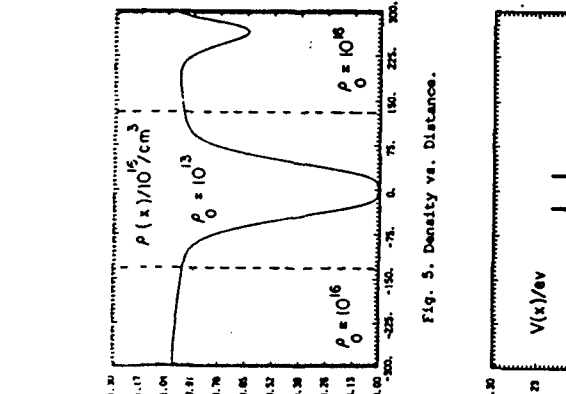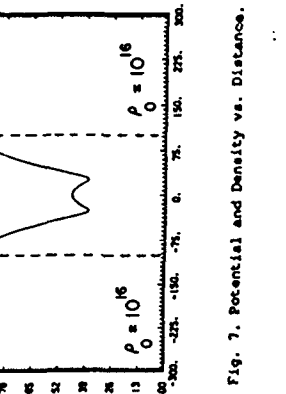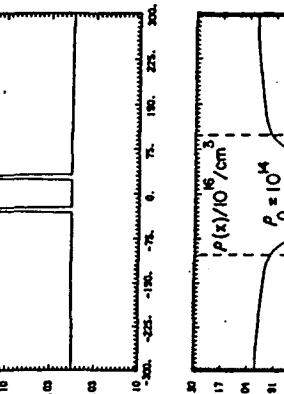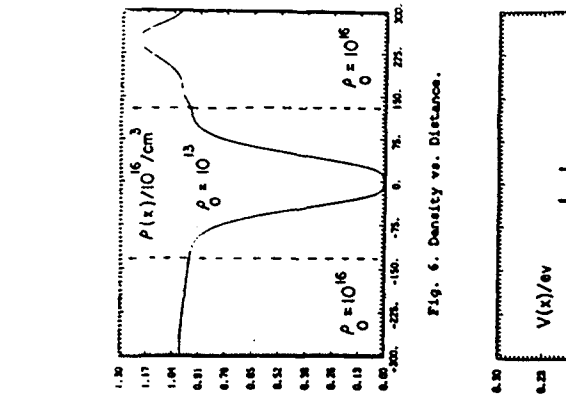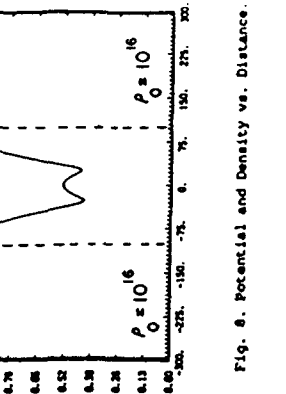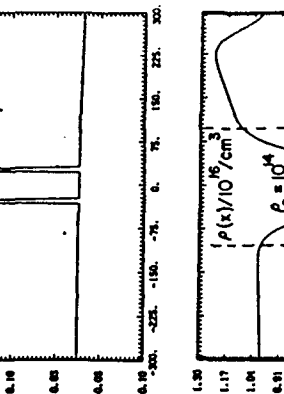Fig. 3. Potential and Density vs. Distance.

Temperature description of transport in single- and multiple-barrier structures

H L Grubin†, T R Govindan†, B J Morrison†, D K Ferry‡ and
M A Stroscio§

†Scientific Research Associates, Inc. Glastonbury, CT 06033, USA
‡Arizona State University, Tempe, AZ 85287-6206, USA
§Army Research Office, Research Triangle Park, NC 27709-2211, USA

Abstract. Barrier calculations based upon solutions of the Liouville equation in the coordinate representation reveal a complicated spatial dependence of the quantum distribution function near and within the barriers. Within the framework of classical transport this spatial dependence suggests equilibrium electron temperature values that differ from the ambient. The prospect of quantum heating and cooling under equilibrium conditions is examined and dispelled in favour of an interpretation that includes density-gradient contributions.

1. Introduction

Calculations based upon solutions of the Liouville equation in a density matrix formulation yield a complicated spatial distribution with a mean kinetic energy in equilibrium that differs significantly from the classical result. In particular, where classical physics teaches that the energy per particle is $k_B T/2$ per degree of freedom for a Boltzmann distribution, quantum physics, as pointed out by Wigner [1], teaches otherwise. The origin of this difference lies in the presence of quantum mechanical forces arising from gradients in density (Ancona and Iafrate, [2]), and are suggestive of a spatially dependent local temperature in both equilibrium and non-equilibrium cases, although spatial dependent carrier temperature in equilibrium introduces interpretive difficulties. To avoid this difficulty one either abandons the spatial-dependent temperature description, or retains it for non-equilibrium studies and seeks another description for equilibrium. But in either case, it is necessary to demonstrate its origin. This is provided below for equilibrium conditions.

2. Energy and temperature

Classical physics indicates that the mean kinetic energy in equilibrium for carriers obeying Boltzmann statistics is

$$\langle E \rangle = \frac{2}{(2\pi)^3} \int d^3 p (p^2/2m) f(x, p) = \tfrac{3}{2} \rho k_B T \qquad (1)$$

$$\equiv \rho \varepsilon$$

where $f(x, p)$ is the classical distribution function, and ε, the mean kinetic energy per particle, is *independent of*

position, as is the consequent electron temperature. For quantum structures, in which quantum distribution functions are required, ε is generally *spatially dependent* [1], and on the basis of equation (1) suggests a spatially dependent equilibrium carrier temperature. Because of the significance of carrier temperature in interpreting hot-carrier phenomena, the spatial dependence of the mean energy per particle, and the origin of this dependence is discussed through solutions to the Liouville equation

$$i\hbar \partial \rho_{op}/\partial t = [H, \rho_{op}] \qquad (2)$$

which in the coordinate representation is a differential equation for $\rho(x, x', t)$

$$\partial \rho/\partial t + (\hbar/2mi)(\nabla^2 - \nabla'^2)\rho$$
$$- (1/i\hbar)[V(x, t) - V(x', t)]\rho = 0. \qquad (3)$$

To expose the essential features of this discussion, we assume Boltzmann statistics, spatial variations only along the x direction, and free particle behaviour along the y and z directions. Transforming equation (3) to centre of mass and non-local coordinates, $r=(x+x')/2$, $\zeta=(x-x')/2$, $\rho \Rightarrow \rho(r+\zeta, r-\zeta)$, we find

$$\rho_t + (\hbar/2mi)\rho_{r\zeta} - (1/i\hbar)[V(r+\zeta, t) - V(r-\zeta, t)]\rho = 0. \qquad (4)$$

In the above equation subscripts denote differentiation. The potential V in equations (3) and (4) is the sum of all heterostructure contributions, $V_0(x)$, and contributions from Poisson's equation:

$$\partial/\partial x[\varepsilon(x)\partial V_{sc}/\partial x] = -e^2[\rho(x, x) - \rho_0(x)] \qquad (5)$$

where the subscript 'sc' denotes self-consistent; $\rho_0(x)$ is the background 'jellium' doping distribution. The diagonal components of solutions to equation (4) (along the diagonal $r = x$ and $\zeta = 0$) provide the density, while the

expectation value of energy $\langle E \rangle$ is obtained from the diagonal components of the kinetic energy density matrix [3]

$$E(r + \zeta, r - \zeta) = -\hbar^2/8m\rho_{\zeta\zeta}. \tag{6}$$

An approximate form of the expectation values of the density and the energy density [1, 4, 5] for one degree of freedom is:

$$\rho(x) = \rho(x, x) = \rho_0 \exp[-(V + Q/3)/k_B T] \tag{7}$$

$$E(x) = [(k_B T/2) + (\hbar^2/24mk_B T)V_{xx}]\rho(x). \tag{8}$$

In equation (7), ρ_0 is a reference density, and $Q(x)$ is the Bohm quantum potential (see, e.g., [6]):

$$Q(x) = -(\hbar^2/2m)(\rho^{1/2})_{xx}/\rho^{1/2}. \tag{9}$$

The second term of equation (8) is referred to as the Wigner contribution. In equilibrium the spatial dependence of the energy per particle, ε, as given by equation (8) is second order in $\hbar$. To this order, if the potential appearing in equation (8) is represented by the Boltzmann relation between density and potential energy, $\rho(x) = \rho_0 \exp[-V(x)/k_B T]$, it is seen that the spatial dependence of ε is a direct consequence of the spatial derivatives of density. In this context the origin of the quantum correction to ε is the same as the origin of the quantum potential.

3. Calculations

The spatial dependence of ε and the origin of the quantum contributions to transport arise from gradients in the carrier density. These features are illustrated through solutions to the Liouville equation for two equilibrium solutions using Maxwellian boundary condition as discussed in [7]. Two cases are considered. For the first calculation a single barrier characterized by a potential

$$V_0(x) = 300(\text{meV})\exp[-(x/12.5\,\text{Å})^2] \tag{10}$$

is placed within a uniform, 1500 Å long, structure doped to $10^{18}\,\text{cm}^{-3}$. The two-dimensional density matrix, $\rho(x, x')$ as obtained from the Liouville and Poisson equations is displayed in figure 1(a). In equilibrium the density matrix is real and symmetric, $\rho(x, x') = \rho(x', x)$, and the solution is completely represented by one-half of the matrix on either side of the diagonal, $x = x'$, as displayed in figure 1(a). The charge density $\rho(x) = \rho(x, x)$ is displayed as a line plot in figure 1(b), where since most of the structure in the solution is contained within a range of 250 Å, about the centre, only 500 Å of the results are displayed. Figure 1(b) displays a significant reduction of charge within the barrier, as well as charge accumulation on either side of the barrier. While the reduction of charge within the barrier is a consequence of the presence of the barrier, the excess charge adjacent to the barrier is a consequence of both self-consistency in the calculation and wavefunction (or density matrix) continuity across the barrier. The spatial dependence of the charge is

consistent with the condition of global charge neutrality. Figure 1(b) also displays two additional plots. The curve reaching the lowest value of density within the barrier is obtained from the classical Boltzmann relation between density and potential energy; the curve reaching intermediate values of density within the barrier is obtained from equation (7). Note that away from the barrier the density from equation (7) approaches a value that is less than the classical value, a result that is a consequence of a change in curvature of the potential as the boundaries are approached. Neither approximate solution can be regarded as an adequate representation of the complete solution, although the quantum-corrected solution possesses the general features of a density that is higher (than classical) within the barrier and lower (than classical) adjacent to the barrier. Figure 1(c) displays the potential distribution. The lowering of the potential adjacent to the barrier ($\sim 20\,\text{meV}$) is a consequence of the excess charge and self-consistency. Figure 1(d) displays the quantum potential; note that its value is greater than $-300\,\text{meV}$ in the centre of the barrier. The energy density matrix represents the curvature of the density matrix in the non-local direction. As seen in figure 1(a), the curvature is steeper where there is excess charge and changes sign within the barrier. The mean kinetic energy per particle, ε, obtained from the density matrix is displayed in figure 1(e), along with the Wigner contribution as obtained from equation (8). It is apparent that the main origin of the structure leading to the Wigner contribution is the quantum potential. The negative value of ε within the barrier and the positive excess value of energy adjacent to the barrier suggest that the Wigner contribution is not a correction, but represents a dominant effect, and that temperature concepts (which must include negative values) are not likely to be germane within the context of equilibrium transport.

The spatial dependence of the mean kinetic energy per particle is also of significance in multiple barrier structures. This is examined for a double barrier structure with

$$V_0(x) = (300\,\text{meV})\{\exp[-(x - 75\,\text{Å})/12.5\,\text{Å}]^2 + \exp[-(x + 75\,\text{Å})/12.5\,\text{Å}]^2\} \tag{11}$$

The barriers are centrally placed within a 1500 Å $n^+n^-n^+$ structure with adjacent $10^{18}\,\text{cm}^{-3}\,n^+$ regions, and a centrally placed 500 Å, $10^{15}\,\text{cm}^{-3}$ region. The two-dimensional density matrix is displayed in figure 2(a), obtained from the Liouville and Poisson equations. There is excess charge between the barriers, a modest increase in curvature between the barriers and a change in sign of the curvature within the barriers. The line plot of density is shown in figure 2(b) over a reduced range of 600 Å. The density as obtained from equation (7) displays a significantly lower charge density within the barrier but *order of magnitude agreement within the quantum well*. The classical solution for density is completely unacceptable. The potential distribution, shown in figure 2(c), reaches flat-band beyond 400 Å on either side of the
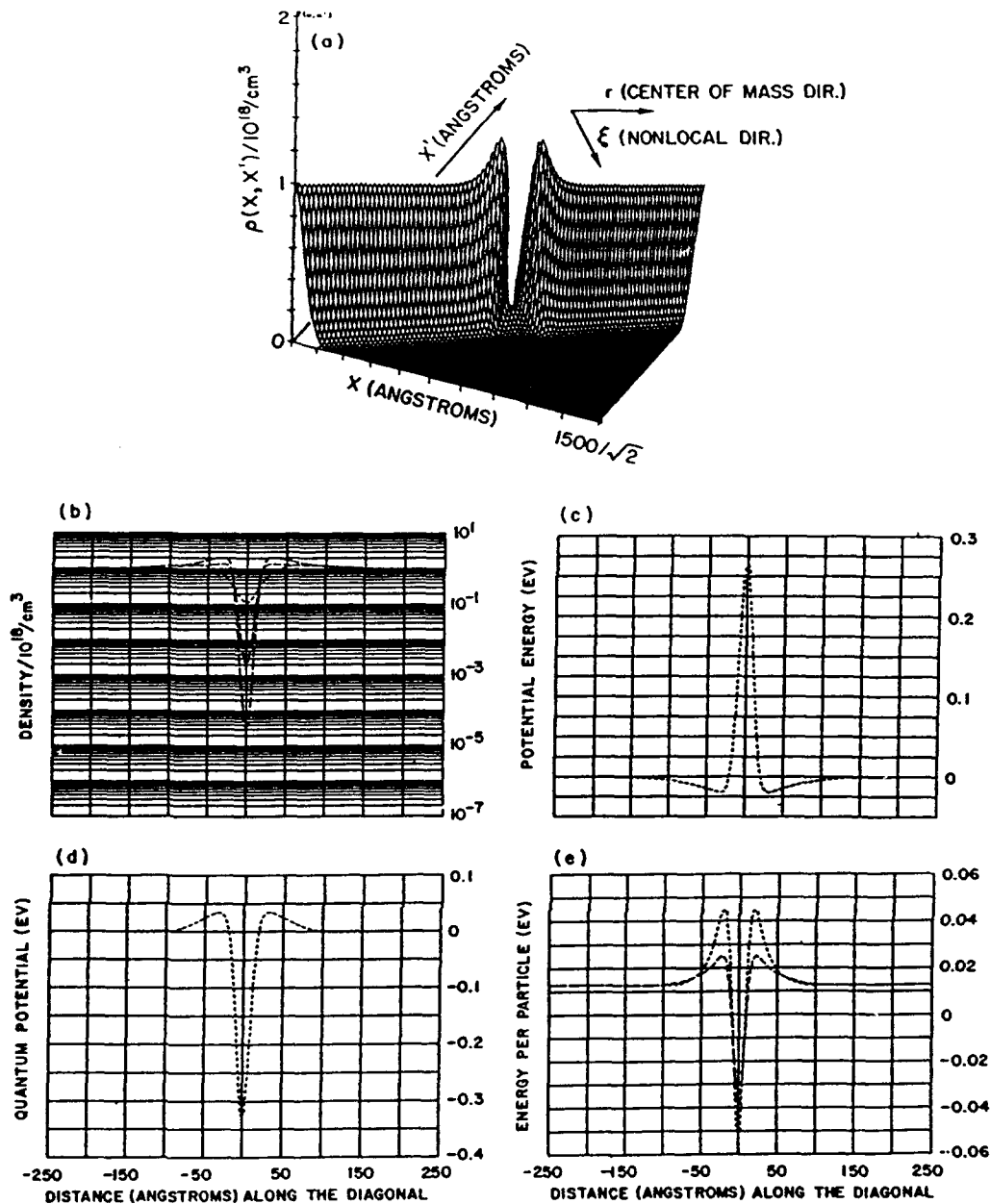
Figure 1. (*a*) Density matrix for a single-barrier structure. The physical dimension of the structure is 1500 Å, requiring that the density matrix, which is calculated over a square matrix, is of side 1500 Å/$\sqrt{2}$. The centre of mass and non-local directions are indicated; (*b*) diagonal component of the density matrix (- - -), from equation (7) (———), classical relation (— — —); (*c*) potential energy $V(x)$; (*d*) quantum potential; (*e*) energy per particle from density matrix (- - -), from equation (8) (— — —).

origin; its increase arises from self-consistency and the reduction of charge in the low-doped region compared with the bounding charge density. The quantum potential displayed in figure 2(*d*) is positive within the quantum well, and emphasizes the reduction in charge density compared with the classical value; it is negative within the barriers, as in the case of the single-barrier structure, and positive outside of the barriers. The positive value outside of the barriers is a consequence of wavefunction and density matrix continuity within the classically accessible region. Note again that the structure of the quantum potential is apparently the main origin of the structure leading to the Wigner contribution to the

energy per particle, as seen in figure 2(*e*). As in the case of the single barrier of figure 1 the calculations suggest that the Wigner contribution is not a correction but represents a dominant effect, and that temperature concepts are not germane in the context of equilibrium transport.

4. Conclusions

The calculations of figures 1 and 2 reveal significant spatial variations in energy associated with density gradients. Mathematically, these energy variations, which are a consequence of wavefunction continuity as repre-
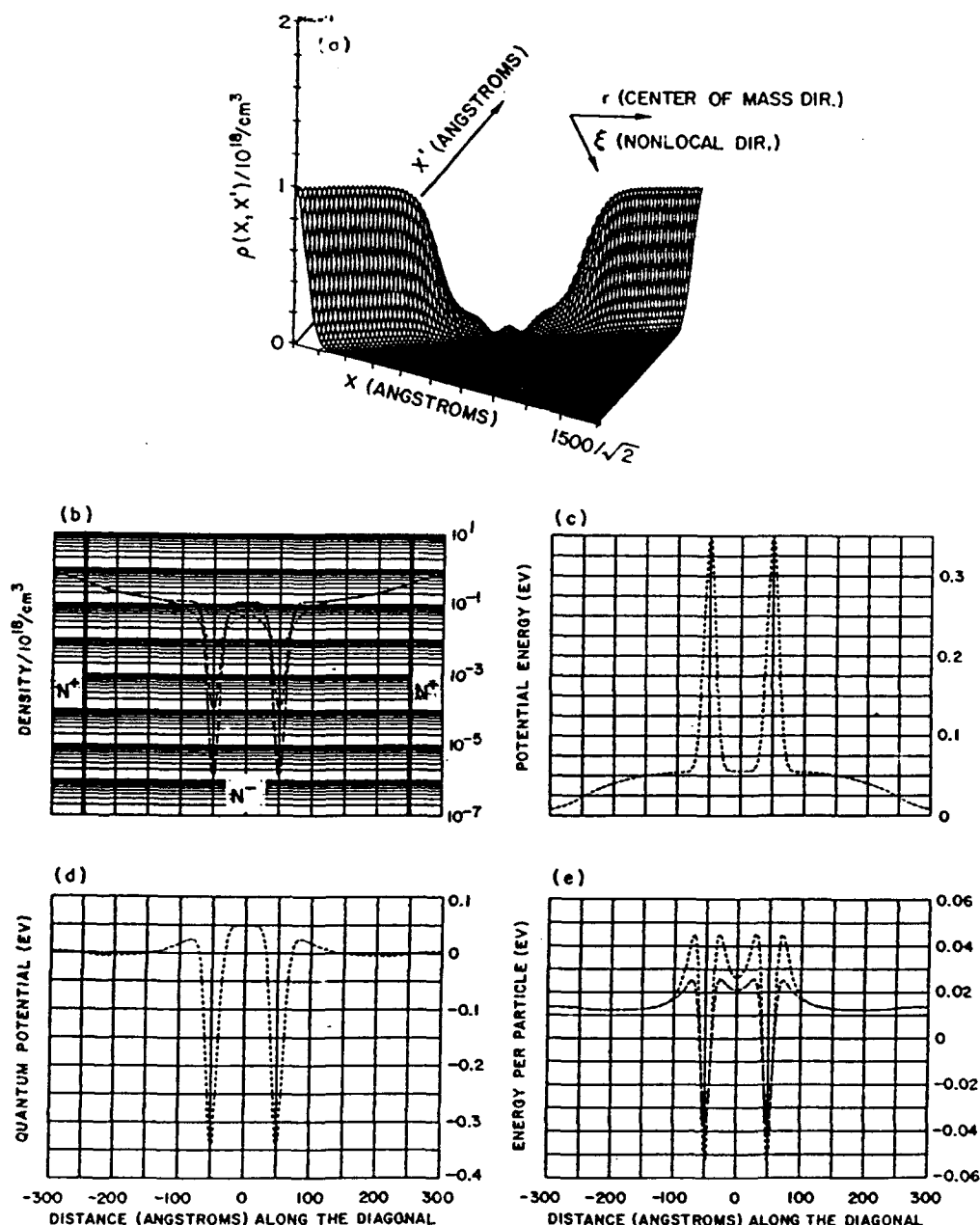
Figure 2. (a) Density matrix for a double-barrier structure; (b) diagonal component of the density matrix (- - -), from equation (7) (———), classical relation (————); (c) potential energy $V(x)$; (d) quantum potential; (e) energy per particle from density matrix (- - -), from equation (8) (———).

sented by 'curvature' in the density matrix, are suggestive of quantum heating and/or cooling. Physically these energy variations represent the influence of local quantum mechanical density dependent forces on the carriers. While their magnitudes indicate that they must be accounted for in all quantum mechanical treatments of transport in mesoscopic structures, an interpretation in terms of heating or cooling in equilibrium is problematic.

Acknowledgement

References

[1] Wigner E P 1932 *Phys. Rev.* **40** 749

[2] Ancona M A and Iafrate G J 1989 *Phys. Rev.* B **39** 9536

[3] Grubin H L, Govindan T R and Stroscio M A to be published

[4] Grubin H L and Kreskovsky J P 1989 *Solid-State Electron.* **32** 1071

[5] Wollard D L, Stroscio M A, Littlejohn M A, Trew R J and Grubin H L 1991 *Proc. Workshop on Computational Electronics*, (Deventer: Kluwer) p 59

[6] Philippidis C, Bohm D and Kaye R D 1982 *Nuovo Cimento* B **71** 75

[7] Govindan T R, Grubin H L and deJong F J 1991 *NASECODE Conference*

Density-matrix and quantum-moment studies of single- and multiple-barrier structures

H L Grubin†, T R Govindan†, B J Morrison† and M A Stroscio‡

†Scientific Research Associates, Inc., Glastonbury, CT 06033, USA
‡Army Research Office, Research Triangle Park, NC 27709-2211, USA

Abstract. The time-dependent Liouville equation for the density matrix in the coordinate representation, incorporating scattering *effects* through a quasi-Fermi level, and Poisson's equation have been solved numerically for a single- and double-barrier structure using algorithms based on 'characteristics', and show significant charge accumulation on the emitter side of the barrier, as well as significant charge screening in multiple-barrier structures.

1. Introduction

Barrier calculations based upon solutions to the Liouville equation with a density matrix formulation display significant charge screening. In particular, the Liouville equation which in the coordinate representation and in the absence of scattering is the differential equation

$$\partial\rho/\partial t + (\hbar/2mi)(\nabla^2 - \nabla'^2)\rho - (1/i\hbar)[V(x,t) - V(x',t)]\rho = 0 \tag{1}$$

yields solutions for $\rho(x, x', t)$ that contain significant quantum departures from the classical solution when there are strong gradients in the carrier density. We illustrate these features for Boltzmann statistics, spatial variations only along the x direction, and free particle behaviour along the y and z directions. We also transform equation (1) to centre of mass and non-local coordinates, $r = (x + x')/2$,

$$\zeta = (x - x')/2, \rho \Rightarrow \rho(r + \zeta, r - \zeta)$$

$$\rho_t + (\hbar/2mi)\rho_{r\zeta} - (1/i\hbar)[V(r + \zeta, t) - V(r - \zeta, t)]\rho = 0. \tag{2}$$

In the above equation subscripts denote differentiation, and the potential V includes all heterostructure contributions, $V_0(x)$, as well as contributions from Poisson's equation:

$$(\partial/\partial x)[\varepsilon(x)\partial V_{sc}/\partial x] = -e^2[\rho(x,x) - \rho_0(x)] \tag{3}$$

The subscript 'sc' denotes self-consistent, and $\rho_0(x)$ is the background 'jellium' doping distribution Note: along the diagonal $r = x$ and $\zeta = 0$.

Significant quantum effects which are revealed through numerical solutions are also revealed through approximate solutions when the potential in equation (2) is expanded in a Taylor series and only the first two non-

zero terms are retained

$$V(r + \zeta) - V(r - \zeta) \approx 2\zeta V_x + \zeta^3 V_{xxx}/3 \tag{4}$$

To second order in $\hbar$, a solution to equation (2) is [1]

$$\rho(r + \zeta, r - \zeta) = \rho_0 \exp[-\beta(V + Q/3) \\ + (\zeta/\lambda)^2(1 + \beta\lambda^2 V_{xx}/6)] \tag{5}$$

where $\beta = 1/k_B T$, $\lambda^2 = \beta\hbar^2/2m$ and

$$Q = (\lambda^2/2)(V_{xx} - \beta(V_x)^2/2). \tag{6}$$

With Q regarded as a quantum correction, inserting the Boltzmann relation between density and potential $(\beta V = -\ln\rho)$, Q is transformed to

$$Q(x) = -(\hbar^2/2m)(\rho^{1/2})_{xx}/\rho^{1/2} \tag{7}$$

which is a generalization of the Bohm quantum potential (e.g., see [2]) for a multiparticle system and teaches that screening is significant when there are strong gradients in carrier density.

While the above discussion ignores scattering, when relaxation effects similar to Fokker–Planck dissipation are introduced the diagonal component of density is approximately given by

$$\rho(x) = \rho(x, x) = \rho_0 \exp[-\beta(V(x) - \xi(x) + Q(x)/3)] \tag{8}$$

with a velocity flux density proportional to $-\rho\partial\xi/\partial x$, and suggests that there are situations in which the effect of scattering on the *density* can be *qualitatively* represented by introducing a scattering operator, *part* of which consists of an algebraic contribution, $\xi(x)$. As a rationale, consider an $n^+n^-n^+$ structure in which the heavily doped regions are long enough for the carriers to relax to background. Under bias assume a quasi-Fermi level exists whose value is such that the density is constant at the upstream and downstream boundaries. This assumption is implicit in most device analyses (e.g., see [3]).

with the conbined effects of an applied bias and scattering *elsewhere* in the structure determining the remaining charge distribution.

In the absence of a detailed description of scattering in structures, how may one expect the density to vary? This is explored through the assumption that an algebraic scattering potential, $\xi(x)$, qualitatively represents the effects of scattering on the density within the structure, and approaches the quasi-Fermi level, E_f, within the boundary regions. $\xi(x)$ is taken as constant and equal to the value of E_f at the emitter boundary $E_f(e)$, until the quantum structure is reached. At the end of the quantum structure $\xi(x)$ is set equal to the value of E_f at the collector boundary, $E_f(c)$. The intermediate values within the quantum region are indicated below. In terms of equation (2), incorporating $\xi(x)$ amounts to replacing $V(x)$ with $V(x) - \xi(x)$.

2. Calculations

The Liouville equation was solved for 1500 Å $n^+ n^- n^+$ structures under bias. Displaced Maxwellian conditions are assumed and a flat band occurs at the boundaries. In one case a single barrier 50 Å wide and 300 meV high was centrally located (figure 1 calculations); in the second case two barriers each 50 Å wide and 300 meV high were centrally placed and separated by 50 Å (figure 2 calculations). For the single-barrier calculations $\xi(x)$ was chosen to ensure that the density at the emitter and collector regions was equal to the background of $10^{18}\,\mathrm{cm}^{-3}$. The values of the algebraic scattering potential across the barrier varied linearly from the value $E_f(e)$ to $E_f(c)$. The small field that forms in the heavily doped regions under bias was ignored and the current through the device was not accurately represented. Higher current values that would normally flow were computed from a simple circuit equation and yielded density distributions similar to that obtained when the current contribution was ignored. The real part of the two-dimensional density matrix $\rho(x, x')$ as obtained from the Liouville and Poisson equations is displayed for a bias range $0 \leqslant V_{\mathrm{applied}} \leqslant -300\,\mathrm{meV}$. $\rho(x, x')$ is symmetric and completely represented by one half of the matrix on either side of the diagonal, $x = x'$. The density matrix shows a build-up of charge on the emitter side of the barrier as well as a broad depletion region on the collector side. Linear plots of the density and potential are displayed in figure 1(e) and 1(f) respectively. As the bias increases there is a lowering of the barrier, and a build-up of a 'notch' potential at the emitter side of the barrier, signifying the development of a region of charge accumulation; the collector side of the barrier shows nearly linear variation in potential suggestive of a broad region of charge depletion. Both accumulation and depletion regions for single-barrier structures have been discussed by Eaves *et al* [4]. Increased charge accumulation on the emitter side of the barrier tends to *reduce* the relative change in potential on the emitter side of the

Note the calculations indicate that global charge neutrality occurs (including the double-barrier calculations); i.e. $\int dx(\rho - \rho_0) = 0$. It is important to note that the quantitative value of the charge adjacent to the barrier is dependent upon the value of $\xi(x)$. For example if $\xi(x)$ were suddenly reduced prior to the barrier, the density prior to the barrier would also be reduced. Global charge conservation would require that local accumulation regions would form elsewhere.

For the double-barrier calculations the low doped region extends over a distance of 600 Å. Two sets of simulations were performed. In the first case the bias on the collector ranged from $0 \leqslant V_{\mathrm{applied}} \leqslant -300\,\mathrm{meV}$. Here, the intermediate value of $\xi(x)$ was equal to $E_f(e)$ to the beginning of the second barrier and varied linearly to $E_f(c)$ at the end of the second barrier. For the second case the bias on the collector ranged from $-100\,\mathrm{meV} \leqslant V_{\mathrm{applied}} \leqslant -500\,\mathrm{meV}$ and the linear variation of $\xi(x)$ started at the beginning of the emitter barrier and ended at the end of the collector barrier. Figures 2(a–d) represent the low-bias range, while figures 2(e–h) are for the high-bias range. Figures 2(a, b and c) display density, potential and $Q + V$ (within select regions of the structure). Figure 2(d) displays the real part of the density matrix for an applied voltage of $-300\,\mathrm{meV}$. The potential variation shows the presence of the formation of a notch region at the emitter side of the first barrier, a smaller fraction of potential falling across the emitter barrier, compared to the collector barrier, as the bias is increased. Indeed, further increases in bias for this variation of $\xi(x)$ result in small variations in potential drop across the emitter barrier region [5]. The density distribution shows the build up of charge in the quantum well, and an increased accumulation on the emitter side of the barrier. At higher bias levels the charge in the well begins to screen the emitter barrier from the collector and the potential drop across the emitter region is smaller than the corresponding case of the single barrier. The term $Q + V$ within the quantum well is displayed because its value is approximately equal to the value of the quasi-bound state energy and moves down as the bias is increased. Successive increases in bias, reduce the movement of the emitter barrier and delay or inhibit the movement of the quasi-bound state to values below the $E_f(e)$, particularly when the n^- regions are very narrow.

The reduction of charge in the well occurs with the variation in $\xi(x)$ as indicated for the second range of bias levels in figures 2(e–h). (Figure 2(h) is the real part of the density matrix for a value of applied potential equal to $-300\,\mathrm{meV}$ and is directly compared to figure (d)). The result is a distribution of charge and potential in the vicinity of the emitter barrier more closely related to that across the single-barrier structure. There is a more equitable potential drop across the double barriers [5], and a skewing of the $Q + V$ terms. A comparison of the potential distribution and charge distribution at $V_{\mathrm{applied}} = -300\,\mathrm{meV}$ indicates that a larger potential drop falls across the double barrier in the absence of charge in the well, which as a consequence will lead to differences in current values and suggests that hysteresis
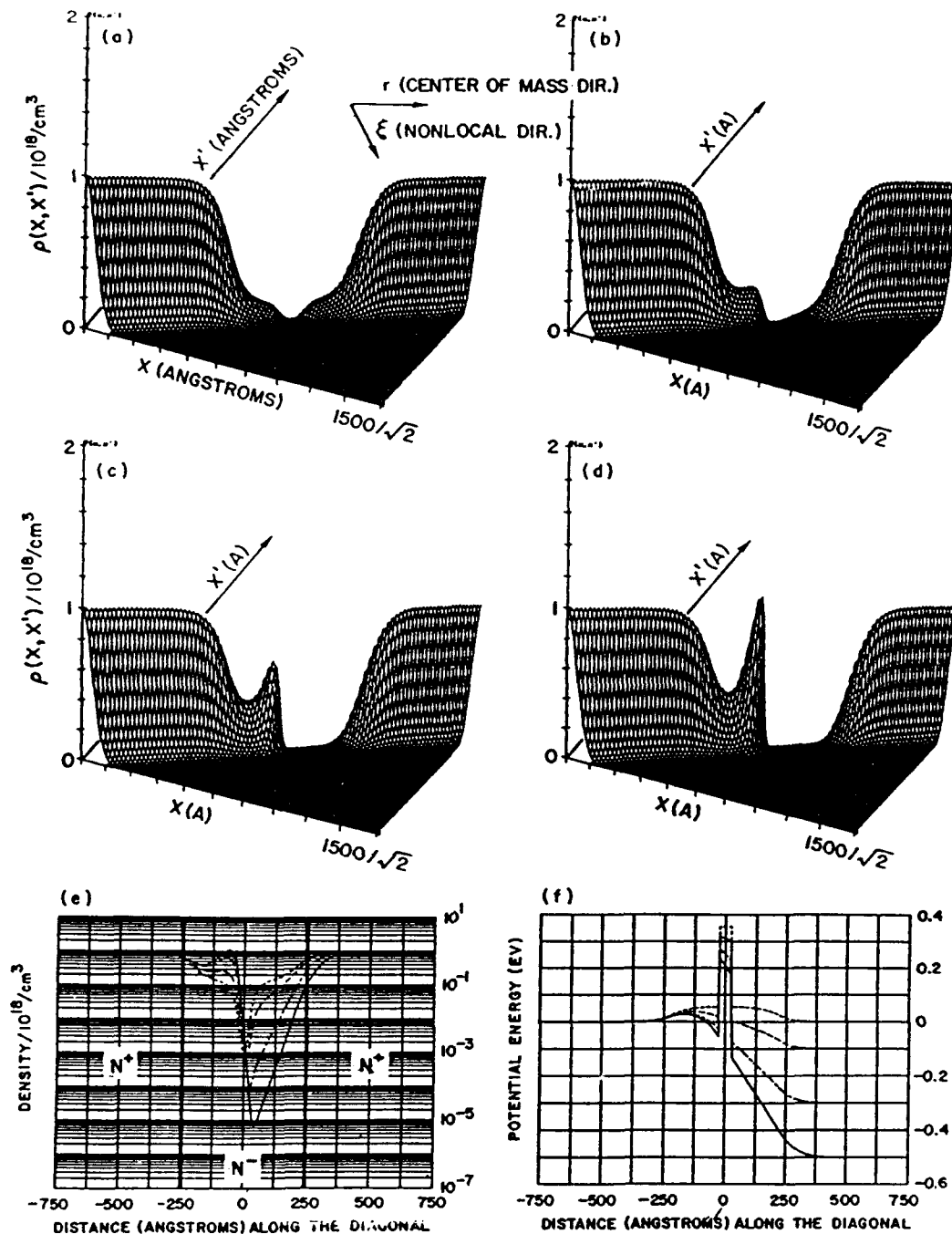
Figure 1. (*a–d*) Density matrix for a single-barrier structure. The physical dimension of the structure is 1500 Å, requiring that the density matrix, which is calculated over square matrix, is of side 1500 Å/$\sqrt{2}$. The centre of mass and non-local directions are indicated. (*a*) $V_{applied} = 0.0$ meV, (*b*) $V_{applied} = -100$ meV, (*c*) $V_{applied} = -300$ meV, (*d*) $V_{applied} = -500$ meV. (*e*) Diagonal component of density matrix: (- - -) 0.0 meV; (———) −100 meV; (— — — —) −300 meV; (——) −500 meV. (*f*) Potential energy (- - -) 0.0 meV; (— — — —) −100 meV; (— · —) −300 meV; (——) −500 meV.

in the measured current voltage characteristics (see, e.g., [4]) is a measure of the variations in the charge distribution as a function of scattering within the structure. Indeed if we assume that the range of scattering potential where $\zeta(x)$ is constant to the collector barrier is representative of the scattering dynamics until the peak current is reached in resonant tunnelling structures, and that scat-

tering results in a reduction of $\zeta(x)$ prior to the collector barrier, for the next and higher range of applied bias levels, then the build up of charge in the well of the resonant tunnelling diode, and the subsequent redistribution of charge to a region upstream of the emitter barrier at higher bias levels emerges from this description.
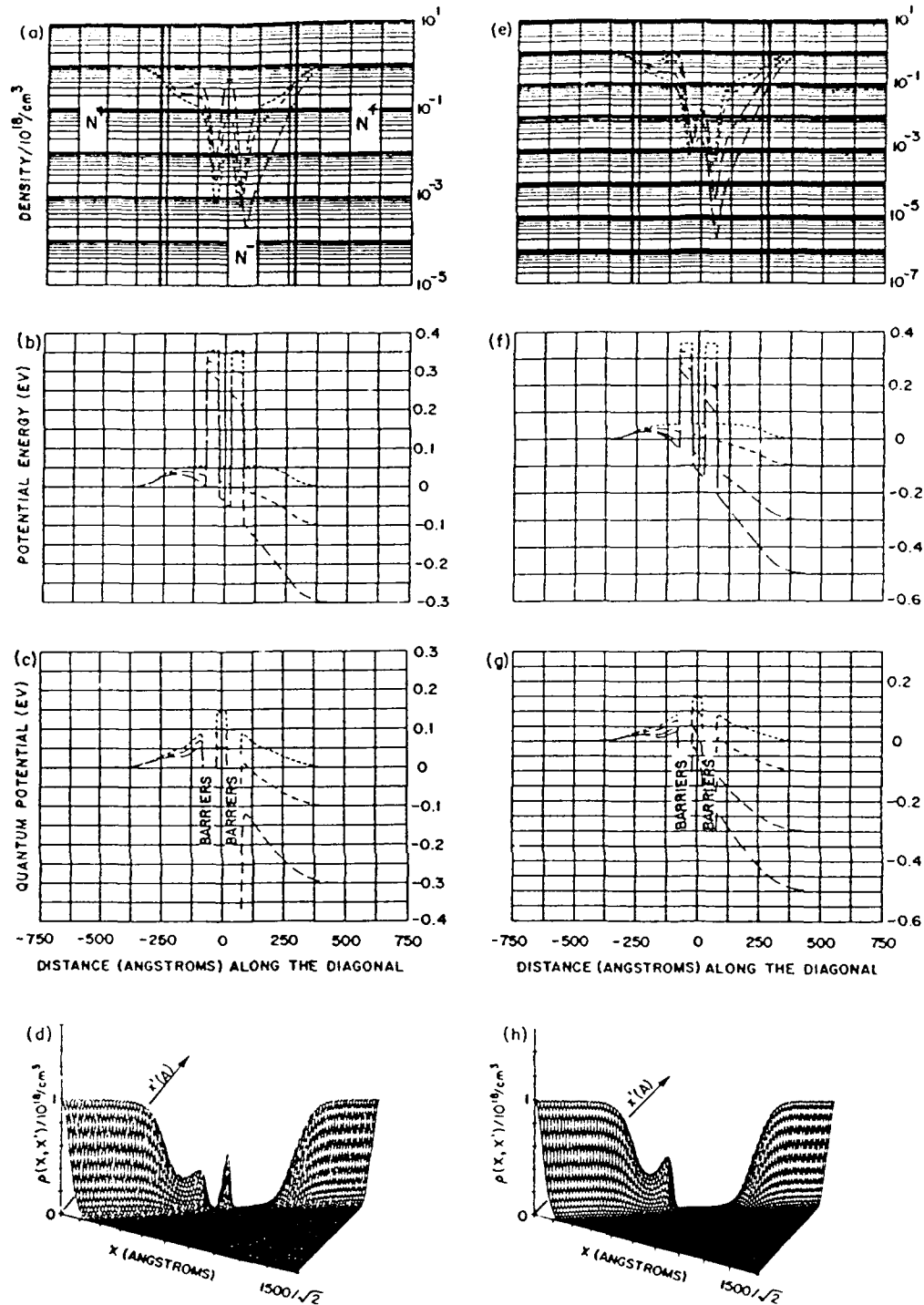
Figure 2. Double-barrier structure under bias: (a and d) density, (b and f) potential energy (c and g) $Q + V$ at select portions of the structure. (d and h) density matrix at -300 meV. Differences in result depend upon algebraic scattering potential (see text). (a), (b) and (c) (- - -) 0.0 meV; (- - -) -100 meV; (————) -300 meV. (e), (f) and (g) (- - -) 0.0 meV; (- - -) -300 meV; (—————) -500 meV.

3. Summary

An appropriate description of transport requires, at least at the boundaries, that the carriers approach their anticipated background values. This is achieved through the introduction of scattering events. The presence of scattering events alters the Liouville equation, with a

consequent change in the distribution of density and in current. In the absence of a detailed description of scattering its effect has been introduced through the incorporation of an algebraic scattering potential. While the results presented here are dependent upon the specific representation of $\xi(x)$ they are representative of the fact the fact that a simple picture of transport as provided by

standard kinetic and potential energy concepts, albeit quantum mechanical, is not likely to be adequate to describe transport in multiple-barrier systems.

References

[1] Grubin H L, Govindan T R and Stroscio M A to be published
[2] Philippidis C, Bohm D and Kaye R D 1982 *Nuovo Cimento* **71B** 75
[3] Shockley W 1950 *Electrons and Holes in Semiconductors* (Princeton, NJ: Van Nostrand Reinhold)
[4] Eaves L, Sheard F W and Toombs G A 1990 *Physics of Quantum Electron Devices* ed F Capasso (Berlin: Springer)
[5] Ricco B and Azbel M Y 1984 *Phys. Rev.* B **29** 1970

1 Introduction to the Physics of Gallium Arsenide Devices

HAROLD L. GRUBIN

Scientific Research Associates, Inc., Glastonbury, Connecticut

1.1 INTRODUCTION

It is arguable that the history of gallium arsenide semiconductor devices, from the early 1960s to the present time, falls into three groups. First, there was the experimental work of Gunn [1], demonstrating the generation of sustained oscillations upon application of a sufficiently large dc bias. This work opened up the possibility of fabricating bulk microwave and millimeter-wave devices, and hastened additional and intense studies of the properties of compound semiconductor devices. Second, there was the study of Ruch [2], whose results suggested that the transient, or nonsteady-state, aspects of semiconductor transport would improve the speed of devices by almost an order of magnitude. This, of course, is the argument behind much of the move toward submicron and ultrasubmicron structures. The third era, the one we are presently in, involves the incorporation of gallium arsenide into material-engineered highly complex structures, some of which have provided remarkable millimeter wave characteristics, such as the pseudomorphic HEMT [3]. Much of this book is concerned with this third era, and thus this chapter will only briefly touch upon it. Rather, this section will present a road map of the consequences of using compound semiconductors for device applications, using gallium arsenide as the paradigm example.

The band structure of gallium arsenide is familiar to most and is displayed in Fig. 1. [4]. It is a direct bandgap material. The minimum in the conduction band is at Γ with relevant subsidiary conduction band minima at L and X. The curvature at Γ is such that the effective mass of the Γ-valley is lower than that of the next two adjacent subsidiary L- and X-valleys. For the valence band, the two valleys of significance are those associated with the light and heavy holes. We will concentrate on transport contributions from these five valleys.

In equilibrium, the relative population of electrons in the valleys is dependent on the density of available states and the energy separation, for
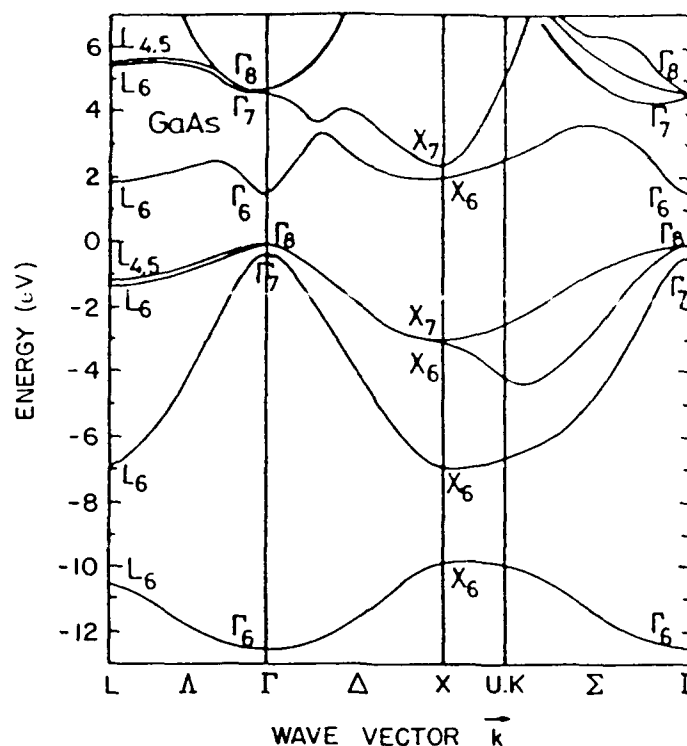
1

Figure 1.1 Band structure of the semiconductor gallium arsenide [4]. (Reprinted, with permission, from *Physical Review* [*Section*] *B: Solid State.*)

example:

$$n_\Gamma^0 = n_L^0 \left(\frac{m_\Gamma}{m_L} \right)^{3/2} \frac{\exp \Delta}{kT} \tag{1}$$

where n_Γ and n_L denote the equilibrium density of the Γ- and L-valley carriers, respectively, m_Γ and m_L are their effective masses, and Δ is the Γ–L energy separation. Thus, in equilibrium virtually all of the electrons of interest are in the Γ-valley. For the holes, the valleys are degenerate.

Gallium arsenide is a compound semiconductor. At low values of electric field, apart from carrier–carrier scattering, there are three important scattering mechanisms: polar optical phonon scattering, acoustic phonon scattering, and impurity scattering. For Γ-valley electrons, the contribution to the momentum scattering rate from polar optical phonons is approximately two orders of magnitude larger than that of the acoustic phonon. Since, with regard to mobility, scattering rates are additive, the polar optical phonon is the dominant scatterer. Ideal room-temperature electron mobilities are in the range of 8000 to 9000 cm^2/V·s. For the subsidiary valleys, the effective masses of the carriers are much larger than that of the Γ-valley, and the relative contribution of the acoustic phonon increases. Nevertheless, the polar phonon dominates the transport. For holes, the situation is mixed,

with the dominant scattering being polar and nonpolar deformation potential coupling. For momentum scattering, the nonpolar deformation potential scattering dominates.

At high values of electric field and for electrons, nonpolar phonons enter the picture, intervalley transfer from Γ to L takes place, and the situation becomes complex. For example, the spatially uniform, field-dependent velocity characteristics of gallium arsenide, ignoring electron-hole interaction, display a region of negative differential mobility, as shown in Fig. 1.2 [5], where at values of field in excess of 3 kV/cm the mean carrier velocity begins to decrease with increasing electric field. This is an unusual situation and it is perhaps important to recognize that the mean electron velocity of a given species of carrier, assuming a parabolic band, is not decreasing with increasing electric field. Rather, the numbers of high-mobility electrons are decreasing, due to transfer to the subsidiary larger effective mass valleys.

The situation with holes is different. Here, the dominant transport is through the heavy hole. Interband hole scattering is always present even at very low fields, however the relative population is fixed through the ratio of the effective masses, and the existence of a dc negative conductance for holes, on the basis of available data, is ruled out. The field dependence of the mean hole velocity, ignoring interaction with the electrons, is displayed in Fig. 1.3 [6], and there are two important features of note. First, there is the extremely low mobility of the holes at low-field values. Second, there is the saturated drift velocity, which is expected to be higher than that of electrons at high fields. We note there is no hard data on the high-field carrier velocity of holes in gallium arsenide.

Calculations of the type displayed in Figs. 1.2 and 1.3 have been described by many workers and are routinely incorporated into simulation codes. Of more recent interest, because of mixed conduction heterostructure devices, are the modifications that may be expected when electron-hole
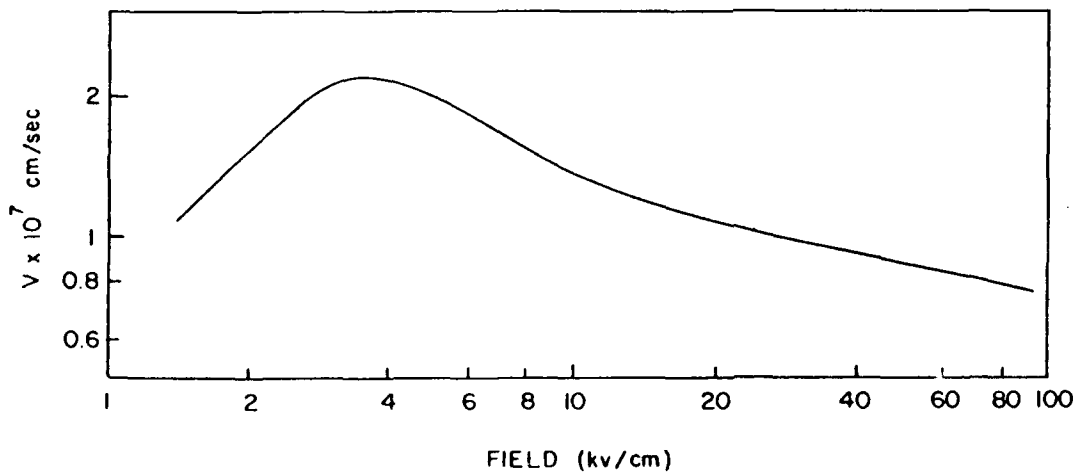

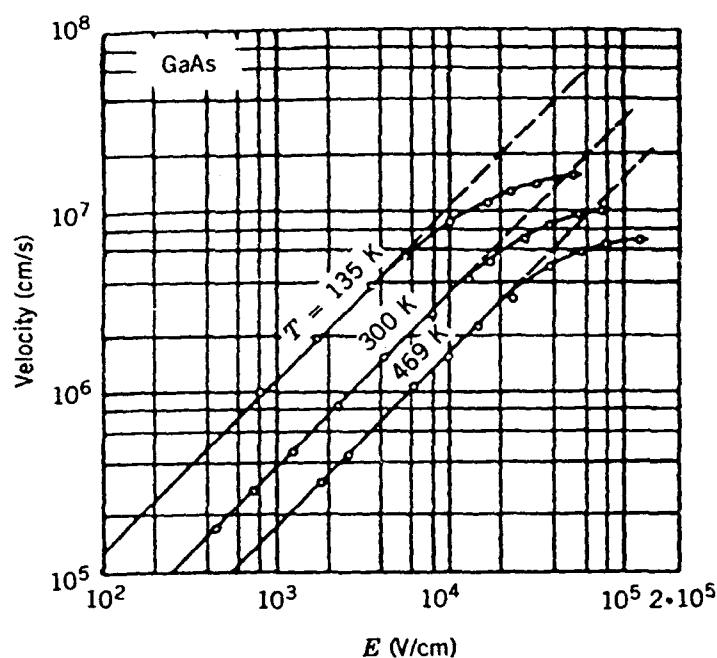Figure 1.2   Field-dependent electron mean velocity for gallium arsenide [5].

Figure 1.3 Field-dependent hole mean velocity for gallium arsenide [6]. (Reprinted, with permission, from *Journal of Applied Physics.*)
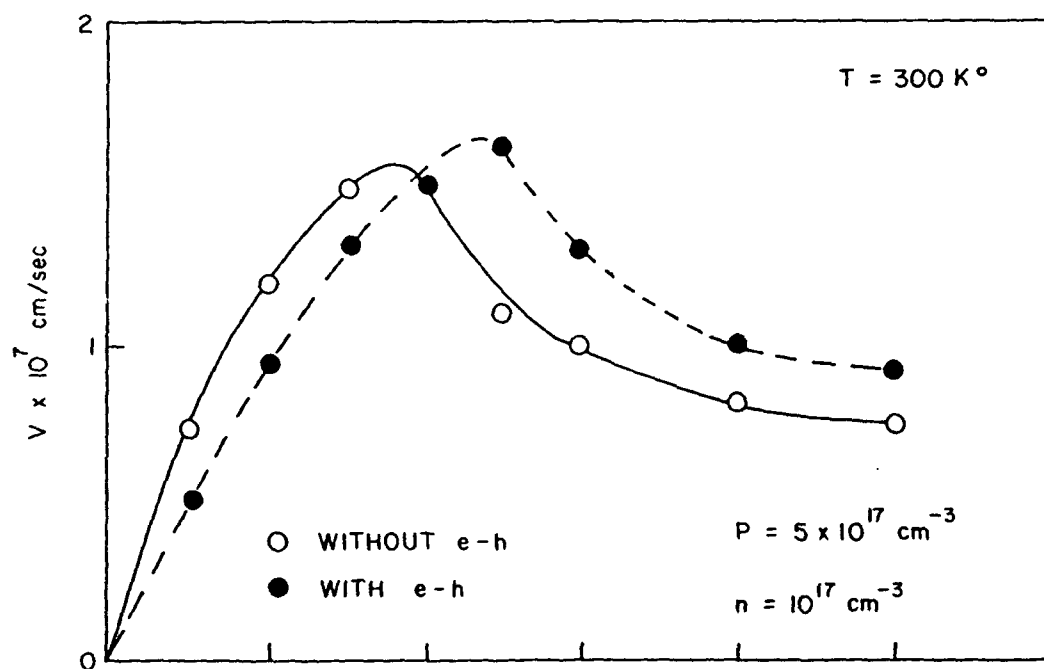

Figure 1.4 Field-dependent electron mean velocity for gallium indium arsenide assuming an interaction with heavy holes [8]. (Reprinted, with permission, from *Applied Physics Letters.*)

scattering occurs. However, because of the limited number of studies with GaAs and because of similarities with other compound semiconductors, results of InGaAs studies are presented. Additionally, because of experimental work in InGaAs [7], the role of carrier–carrier scattering has been most extensively studied for that material. Monte Carlo calculations incorporating electron-hole scattering are displayed in Fig. 1.4 [8]. The results require some detailed discussion and are considered later. Here, we simply note that the presence of holes leads to reduction in the low-field mobility but an increased peak carrier velocity. These intriguing results are also anticipated for GaAs.

The question of interest is how may we expect the role of the complicated compound semiconductor band structure to affect the performance of devices. This is considered next.

1.2 THE ROLE OF BAND STRUCTURE ON THE OPERATION OF ELECTRON DEVICES

In examining the role of band structure on the operation of electron devices, there are several items of immediate interest: the effective mass, the low-field mobility, and the direct bandgap energy of the binary III–V materials (see Table 1.1 [9]). Additionally, the energy separation of the conduction band minima to the subsidiary valleys is listed in Table 1.2 [10]. Note that of the seven binary materials listed, five are direct bandgap materials, and two, GaP and AlAs, are indirect materials. The indirect

TABLE 1.1   Critical Parameters of Select Compound Semiconductors

Compound	Effective Mass[a]	Electron Low-field Mobility $(cm^2/V \cdot s)$	Direct Energy Bandgap (eV)
GaAs	0.063^b	9,200	1.424
GaP[c]	$0.25^i/0.91^i$	160	2.78
GaSb	0.042	3,750	0.75
InAs	0.0219	33,000	0.354
InP	0.079	5,370	1.344
InSb	0.0136	77,000	0.230
AlAs[d]	0.71^b	300	2.98

Source: Ref. 9. Reprinted with permission of Springer-Verlag.
[a]Multiples of free electron mass at the conduction band minima.
[b]DOS.
[c]The minima in the conduction band are at Δ-axis near zone boundary.
[d]The minimum in the conduction band is at X.

TABLE 1.2   Intervalley Energy Separation

Compound	$\Gamma-L$ (eV)	$\Gamma-X$ (eV)	$L-X$ (eV)
GaAs	0.34	0.48	0.14
GaP	−0.27	−0.39	−0.37
GaSb	0.08	0.37	0.23
InAs	1.27	1.60	0.33
InP	0.63	0.73	0.10
InSb	0.41	0.97	0.56
AlAs	−0.15	−0.79	−0.64

Source: Ref. 10. Reprinted with permission from Oxford University Press.

bandgap materials have the highest effective masses of the group and also the lowest mobility. Of these materials, GaAs, InAs, InP, and InSb possess regions of negative differential mobility. GaSb, GaP, and AlAs do not. It is perhaps not surprising that the first four mentioned materials possess a region of negative differential mobility, nor that the last two materials do not. In the latter case, the minima in energy is associated with a large effective mass, high density-of-states energy level. The situation with GaSb is peculiar. But here, while the effective mass of the Γ-valley is the smallest of the three, its closeness in energy to that of the subsidiary L-valley is such that at low values of field conduction, contributions arise from both the Γ- and L-valley, effectively suppressing the contributions of intervalley transfer to negative differential conductivity.

The presence of a region of bulk negative differential mobility has, as a major consequence, the possibility of electrical instabilities. These instabilities manifest themselves either as large-signal dipole-dominated oscillations, often referred to as the Gunn effect, or as circuit-controlled oscillations, where the semiconductor behaves electrically as a van der Pol oscillator. The binary semiconductors GaAs, InP, and InAs have exhibited electrical instabilities associated with bulk negative differential mobility. While InSb has also sustained electrical instabilities, the interpretation of the instability is complicated by the small direct bandgap and the possibility of avalanching at low bias levels.

An additional feature of importance is the intrinsic carrier concentrations of some of these materials, as shown in Table 1.3 [9]. It is clear that the intrinsic concentration of InAs and InSb make them unsuitable for a unipolar source. Indeed, all transport calculations using these latter materials must necessarily include multispecies transport.

In choosing materials for electron devices, particularly as power sources, a figure of merit has been the peak to saturated drift velocity ratio. From this point of view, indium phosphide is an attractive candidate, but this must be weighed with the fact that the low-field mobility of InP is less than that of

TABLE 1.3   Intrinsic Concentration

Compound	n ($/cm^2$)
GaAs	2.1×10^6
InAs	1.3×10^{15}
InSb	2.0×10^{16}
InP	1.2×10^8

Source: Ref. 9. Reprinted with permission of Springer-Verlag.

gallium arsenide. A recent study comparing these features suggests that the Γ-valley mobility is the dominant material parameter of submicron structures, whereas the high-field saturated drift velocity is the dominant material parameter of micron-length structures [11].[†] Additionally, if a choice for two terminal sources is to be made between, for example, InP and GaAs, other issues emerge. For instance, the scattering rates in InP indicate a shorter energy relaxation time than that of GaAs. The consequence of this are higher-frequency *operation for InP*. Thus, at least with respect to these materials, the peak to valley ratio of the materials is only one factor in the design of an electrical source.
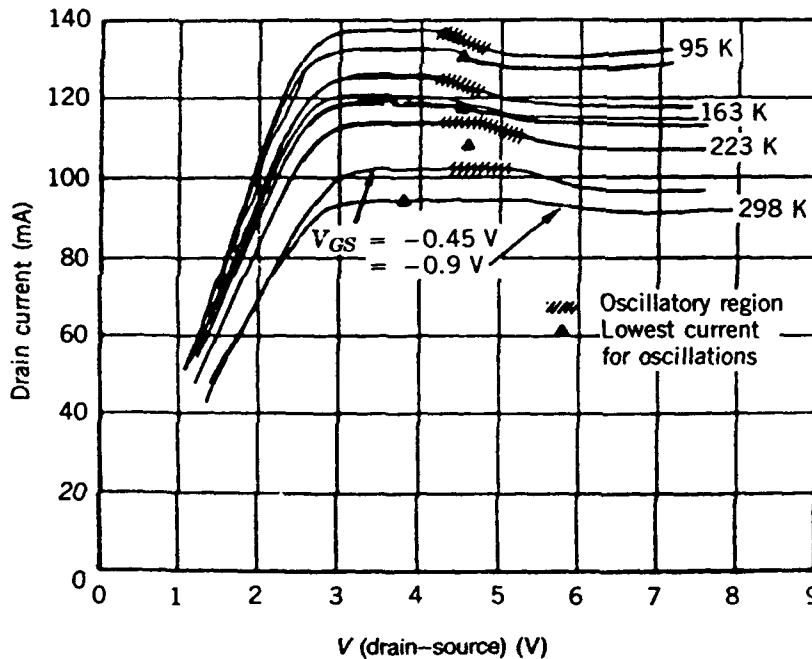

Figure 1.5 Temperature-dependent pulsed data for a GaAs FET, with a 3.0 μm gate length, a source to drain separation of 8.5 μm, and an epitaxial thickness of 3000 ± 500 A. Nominal background doping is $10^{17}/cm^3$ [12].

[†](Editor's Note) See, however, Chap. 2, page 82, last paragraph and reference [12] cited within for an additional view on this issue.

There is less to say about the effects of negative differential mobility on the operation of avalanche diodes. Here, the effects of negative differential mobility conductivity are present but are overshadowed by the effects of avalanching. For example, recent simulation studies show the presence of domains in IMPATTs, whose presence is a direct consequence of negative differential conductivity. These domains can complicate the actual transit time of dipole layers associated with the avalanche generation, but the negative differential mobility is a marginal issue. Such is not the case with three-terminal devices.

For three-terminal device observations of bias-dependent white light in GaAs FETs, as from either the drain side of the gate contact and the gate side of the drain contact, are consistent with numerical calculations showing the presence of local high-field dipole layers near the gate and drain contacts. In addition, for a range of bias, some devices display a current dropback consistent with bias-dependent formation of high-field domains and concurrent current oscillations. This last result is shown in Fig. 1.5 [12]. Remaining questions of interest focus on the manner in which transport in these devices is examined. We begin with the equilibrium description of transport.

1.3   EQUILIBRIUM DESCRIPTION OF TRANSPORT

The steady-state equilibrium description of transport has traditionally provided most details of device behavior. Nevertheless, the description ignores acceleration. It assumes that the carrier velocity is determined by the local electric field and that the total current is governed by a balance of a drift component and a diffusive component. Typically, the continuity equation is solved simultaneously with the current equation, which for electrons is of the form

$$\mathbf{J}_n = -e\left(n\mathbf{v}_n - D_n \frac{\partial n}{\partial \mathbf{x}}\right) \tag{2}$$

and for holes:

$$\mathbf{J}_p = +e\left(p\mathbf{v}_p - D_p \frac{\partial p}{\partial \mathbf{x}}\right) \tag{3}$$

Here, n and p denote electron and hole concentration, respectively, $\mathbf{v}$ velocity, and $\mathbf{D}$ diffusivity. The usual derivations of Eqs. (2) and (3) proceed from a linearization of the Boltzmann transport equation. The assumption is then made that the equation is valid for high-field nonlinear transport.

Typically, the field-dependent velocity assumed in these equations is of the type displayed in Fig. 1.2.

While the use of the field-dependent velocity in these equations is universal, the type of diffusivity coefficient used in these studies is almost as numerous as the numbers of workers involved in numerical studies. However, a number of important issues are at stake in the description of the diffusivity. For example, if the Einstein relation

$$D = \frac{\mu kT}{e} \qquad (4)$$

is used, then, under equilibrium and/or zero current conditions, the dependence of carrier density on conduction and valence band energy is given by either the equilibrium Boltzmann or Fermi distribution. However, under nonequilibrium conditions (and near-zero current conditions), the Einstein relation inadequately describes diffusive transport [13]. To correct for the latter deficiency, the field-dependent diffusivity often used in calculations is of a form similar to that shown below [14]:

$$D = \frac{\mu kT}{e} + \tau v_{sat}^2 \qquad (5)$$

where at high values of electric field, the diffusivity only gradually decreases. While the diffusivity coefficient of Eq. (5) more adequately represents high-field phenomena, because its field dependence is conceptually consistent only with the assumption of nonequilibrium conditions it is conceptually inconsistent with equilibrium conditions, and will lead to incorrect built-in potentials [15].

While the drift and diffusion equations (DDE) clearly offer conceptual difficulties with respect to consistency of physics, they nevertheless offer considerable insight into the physics of device operation and are useful providing their limitations are kept in mind. For example, instabilities in long GaAs structures are known to depend critically on conditions at the contacts. A study in 1969 [16] demonstrated that by experimentally creating different conditions at the boundaries to the active region of GaAs, a wide range of different electrical instabilities could be obtained. Corresponding numerical studies were performed through solutions to the above drift and diffusion equations, in which a value for the electric field was specified at the cathode (and anode) boundary. It was found that the boundary-dependent electrical behavior could be broken into three categories, as summarized in Fig. 1.6 [16]. The key conclusion of the study was that the electrical behavior of compound semiconductors devices was dependent in a detailed way on contact conditions. This same critical result has reappeared numerous times in a variety of different types of structures.
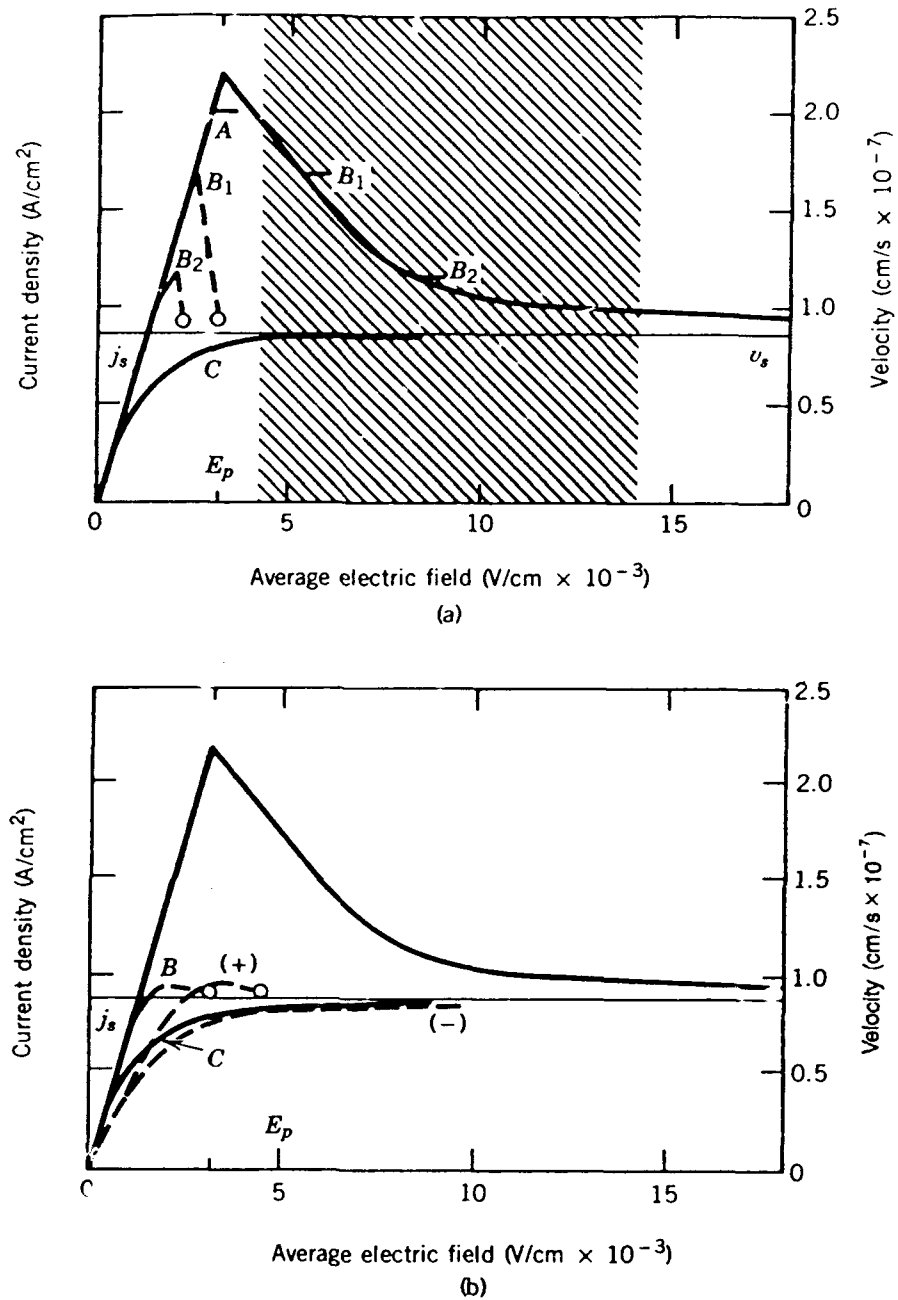
Current density (A/cm^2)

Velocity (cm/s × 10^{-7})

A

B_1

B_1

B_2

B_2

j_s

C

v_s

E_p

Average electric field (V/cm × 10^{-3})

(a)

Current density (A/cm^2)

Velocity (cm/s × 10^{-7})

B (+)

j_s (−)

C

E_p

Average electric field (V/cm × 10^{-3})

(b)

Figure 1.6 (a) The $v(E)$ curve and the computer simulated current density j as a function of average electric field $\langle E \rangle$ (A through C) for various cathode boundary fields. The boundary field is zero for curve A, 24 kv/cm for curve C, and is indicated by the arrow for curves B_1 and B_2. The sample is 10^{-2} cm long and has $n = 10^{15}$ cm^{-3} and $\mu = 6860$ cm^2/V·s. The right- and left-hand ordinates are related by $j = nev$, $v_2 = 0.86 \times 10^7$ cm/s. (b) Experimental $j - \langle E \rangle$ curves (+) and (−) (dashed) and theoretical curves B and C (solid). The only significance in the fact that the low-field slopes differ is that the theoretical curve is for a mobility of 6860 cm^2/V·s, whereas the experimental curve is for a mobility of 4000 cm^2/V·s [16].

1.4 NONEQUILIBRIUM DESCRIPTIONS OF TRANSPORT

The situation of most interest lies in nonequilibrium transport. The most critical area of interest is the incorporation of acceleration into the governing equations.

In examining nonequilibrium transport, several approaches have been used. One is the Monte Carlo method, where the trajectory of a particle is followed through its acceleration and subsequent scattering events. In the discussion below, results of Monte Carlo calculations will be presented, but we first concentrate on nonequilibrium phenomena as described by the moments of the Boltzmann transport equation. These equations, in their simplest form for parabolic bands, a position-dependent conduction band, and a position-dependent effective mass, take the form shown below [17]:

Carrier balance:

$$\frac{\partial n}{\partial t} + \nabla_r \cdot n \frac{\hbar \mathbf{k}_d}{m} = \frac{2}{(2\pi)^3} \int \frac{\partial f}{\partial t}\bigg|_{coll} d^3 \mathbf{k} \qquad (6)$$

Momentum balance:

$$\frac{\partial}{\partial t} n\hbar \mathbf{k}_d + \nabla_r \cdot \frac{(n\hbar^2 \mathbf{k}_\_ \mathbf{k}_\_)}{m} = -n\nabla_r E_c + qn\mathbf{v} \times \mathbf{B} - \nabla_r nkT$$

$$+ \left(n \frac{\hbar^2 \mathbf{k}_d \cdot \mathbf{k}_d}{2m} + \frac{3}{2} nkT \right) \frac{\nabla_r m}{m} + \frac{2}{(2\pi)^3} \int \frac{\partial f}{\partial t}\bigg|_{coll} \hbar \mathbf{k} \, d^3 \mathbf{k} \qquad (7)$$

Energy balance:

$$\frac{\partial}{\partial t} \left[n\left(\frac{\hbar^2 \mathbf{k}_d \cdot \mathbf{k}_d}{2m} + \frac{3}{2} kT \right) \right] + \nabla_r \cdot nv\left(\frac{\hbar^2 \mathbf{k}_d \cdot \mathbf{k}_d}{2m} + \frac{5}{2} kT \right)$$

$$= -nv \cdot \nabla_r E_c + \frac{2}{(2\pi)^3} \int \frac{\partial f}{\partial t}\bigg|_{coll} \frac{\hbar^2 \mathbf{k} \cdot \mathbf{k}}{2m} d^3 \mathbf{k} \qquad (8)$$

In the above, $\hbar \mathbf{k}_d$ is the mean momentum of the carriers, T is the carrier temperature, and, for electrons, E_c is the position-dependent conduction band energy. $\mathbf{B}$ is an applied magnetic field. The terms on the right side represent scattering and/or electron-hole interaction, as through avalanching. For example, the right side of Eq. (6) represents intervalley scattering. If avalanching occurs, generation is expressed through an energy-dependent ionization coefficient [18]. If a carrier temperature model is assumed, then

carrier generation is given by

$$n\alpha(T) \tag{9}$$

where $\alpha(T)$ is the ionization coefficient. In the absence of a first-principle determination of $\alpha(T)$, the following relation can be assumed as a starting point:

$$\alpha(T) = \alpha^*(F)v(F) \tag{10}$$

where $\alpha^*(F)$ and $v(F)$ are the equilibrium ionization rates and field-dependent velocities, respectively and the relation between T and F is determined from the equilibrium solution. While the Eq. (10) relation is uncertain, it has the conceptual advantage of relating ionization to energy, rather than field.

But, perhaps the most significant feature of these equations is the presence of acceleration, both spatial and temporal in the momentum balance equation. These acceleration terms are absent from the drift and diffusion equations. Additionally, under equilibrium conditions, and hence, zero current (i.e., $n\hbar k_d/m = 0$), the electron temperature model teaches that for any spatially nonuniform structure, such as a *p-n* junction, the electron temperature is everywhere constant and equal to the ambient. Thus, co. ceptual problems arising from the form of the diffusion contribution to the drift and diffusion equations do not enter here. Note that a generalized drift and diffusion current term is obtained when the left side of Eq. (7) is set to zero.

Equations of the type shown above provide a considerable amount of information with regard to transport. For example, with a $\Gamma-L-X$ orientation in GaAs the distribution of carriers as a function of field is shown in Fig. 1.7 [5]. Here the relative distribution of carriers in each of the valleys is determined by the distribution of temperature in each of the valleys, which in turn is driven by the electric field, as shown in Fig. 1.8 [5]. Note that for fields below 4 kV/cm, the carriers reside in the Γ-valley. At fields above 6 kV/cm, the L-valley population exceeds that of the Γ-valley. It should be emphasized, however, that because of the very low subsidiary valley mobility, most of the current, for fields up to 50 kV/cm, is carried by the Γ-valley carriers.

The interest in nonequilibrium equations lies not in the steady-state uniform field distribution, but in transients and nonuniform fields. The transient distribution of carrier density and velocity for electrons subject to a sudden change in electric field is shown in Fig. 1.9 [5], where the high peak velocity can be noted.

The high peak velocity in Fig. 1.9a is primarily associated with Γ-valley transport. Indeed, the Γ-valley velocity at 2.8 ps is near 4×10^7 cm/s. The
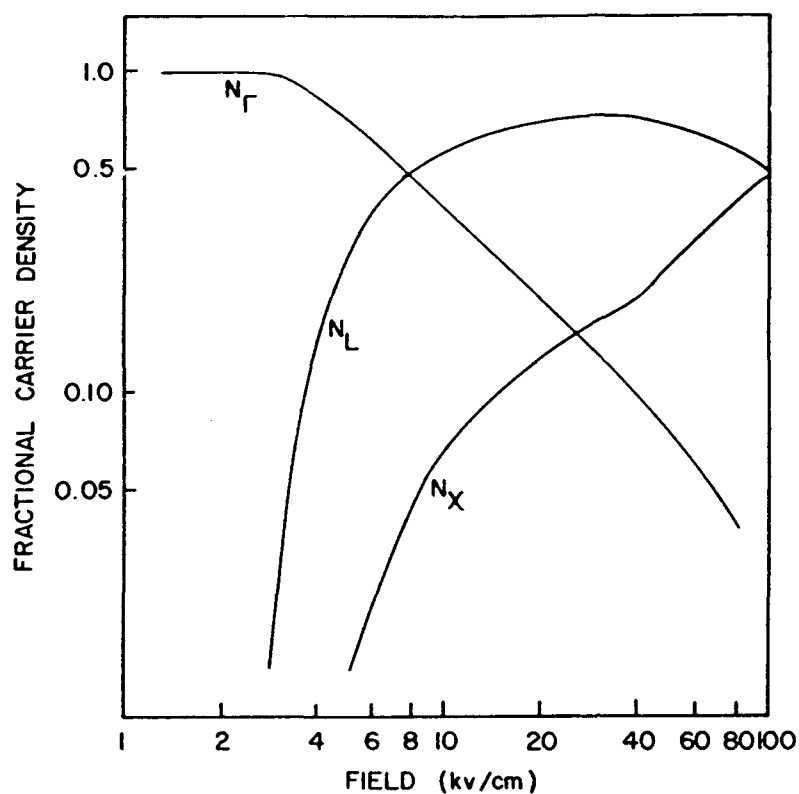
Figure 1.7   Steady-state uniform field carrier distribution for $\Gamma-L-X$ orientation in GaAs [5].
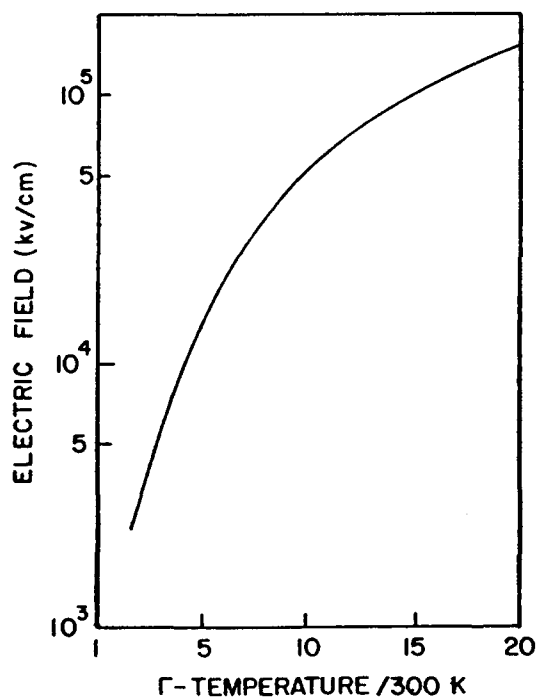

Figure 1.8   Temperature dependence of electric field for a $\Gamma-L-X$ orientation in GaAs. Electric field is the independent variable [5].
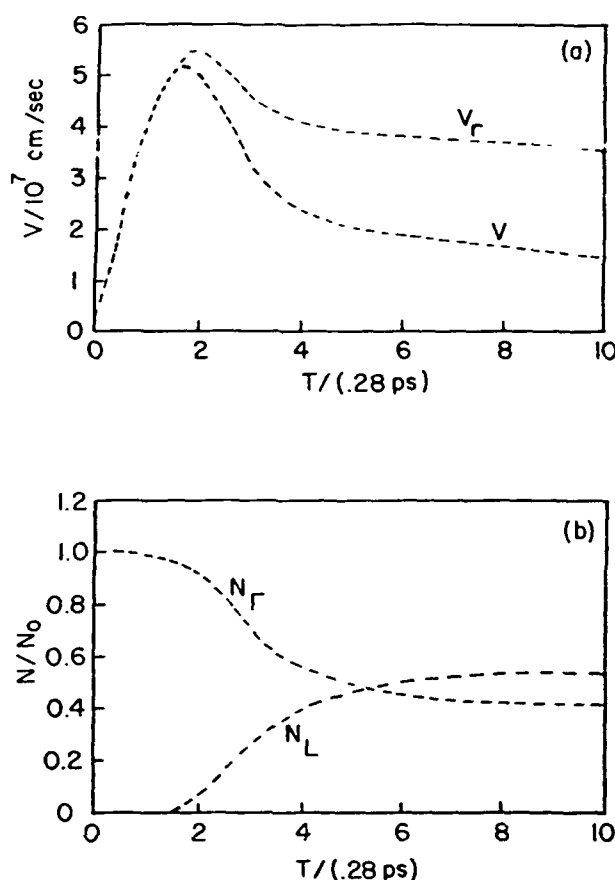
Figure 1.9 Transient overshoot for a field of 9.6 kV/cm [5]. Γ–L–X orientation, and an applied field of (a) Γ and mean velocity, (b) Γ- and L-valley carrier density.

mean velocity

$$v = (n_\Gamma v_\Gamma + n_\Gamma v_L + n_X v_X)/N_0 \qquad (11)$$

where n denotes net population of the Γ-, L-, and X-levels and N_0 is the total carrier density, is also shown in Fig. 1.9. We note that the significant drop in mean velocity is a consequence of electron transfer from the central to the satellite valleys (see Fig. 1.9b). Also note a transient decrease in Γ-valley velocity. This is a consequence of the difference between the energy (longer) and momentum (shorter) relaxation times in GaAs. The time-independent spatially nonuniform situation also displays overshoot effects.

The situation when, in one dimension, space charge effects are introduced is displayed in Figs. 1.10 and 1.11 [5], where for a gallium arsenide device of different lengths we show the field and space charge distribution of the Γ-valley electrons and the current–voltage characteristics. The feature to note is that as the device length decreases the current drive increases. Note that in all cases the field is nonuniform and increases toward the anode. Electron transfer exists for all four structures, with the greatest amount of transfer occurring for the longest device. Additionally, since high-field
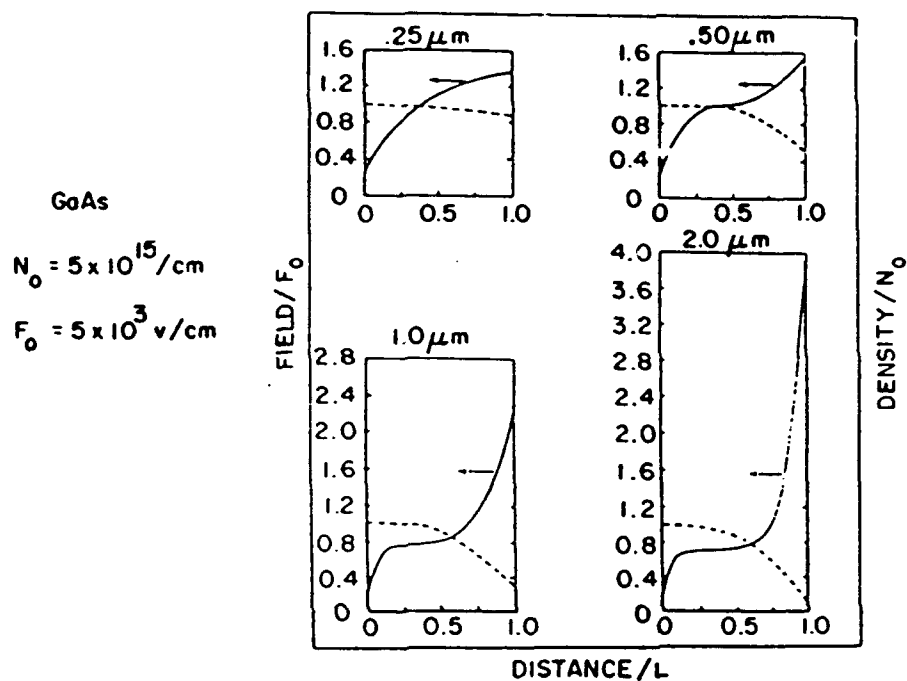
Figure 1.10 Electric field and Γ-valley carrier distribution for GaAs (with a two-level model) two-terminal devices of different lengths and a mean field of 5 kV/cm [5].
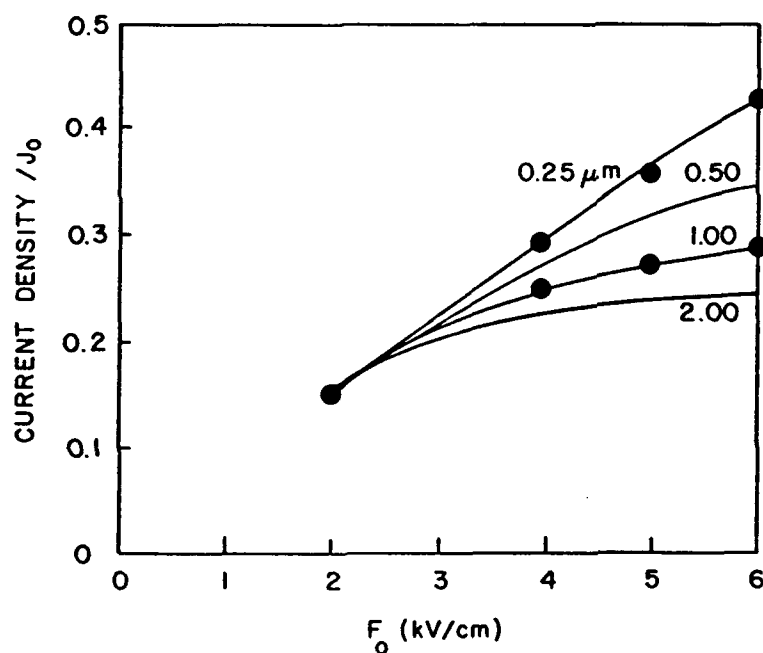

Figure 1.11 Current-voltage characteristics, as a function of lengths, for the structures of Fig. 1.10 [5].
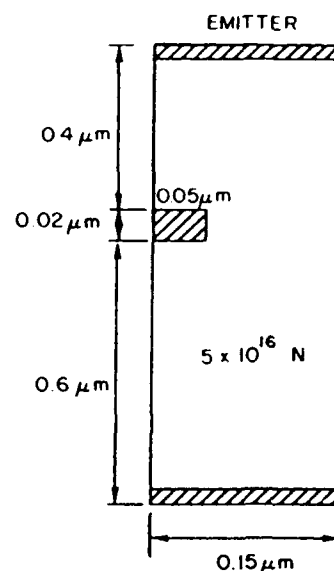
EMITTER

0.4 μm

0.05 μm

0.02 μm

5 × 10^{16} N

0.6 μm

0.15 μm

Figure 1.12 Dimensions and doping level of the simulated PBT [20].

values are synonymous with carrier accumulation, we see that electron transfer here is accompanied by an accumulation of L-valley carriers. Saturation in the current density occurs at high bias, and even for the shortest device there is electron transfer at the anode side of the device.

The role of nonequilibrium transport on two-dimensional simulations is discussed for the vertical three-terminal GaAs permeable base transistor [19], one cell of which is displayed in Fig. 1.12 [20]. One important advantage of this structure is the parallel placement of the source and drain contacts* and the absence of any substrate through which current can flow and reduce the transfer characteristics of the device. The simulations were performed for a 1 μm long source to drain region and a 200 Å gate. Also, for comparison, results of the drift and diffusion equation simulations were included.

The computed I–V characteristics of the device shown in Fig. 1.12 are presented in Fig. 1.13 [20]. The moments of the Boltzmann transport equation results are extrapolated to the origin, as indicated by the long broken lines. The shorter dashed curves show the results for the DDE. The comparison shows that the predicted current levels are significantly higher for the moment equation solutions, a result consistent with FET calculations performed by Cook and Frey [21], who used a highly simplified momentum-energy transport model. The present calculation results also indicate a region of negative differential forward conductivity at $V_{BE} = 0.6$ V. The origin of this phenomena is a consequence of electron transfer.[†] The presence of a dc negative forward conductance is also a feature of PBT measurements but is clearly absent from DDE simulations [20].

*(Editor's Note) They are also called emitter and collector in PBT terminology.

†(Editor's Note) See also Chap. 2, Section 2.2.4 for additional views on the issue of negative differential conductance.
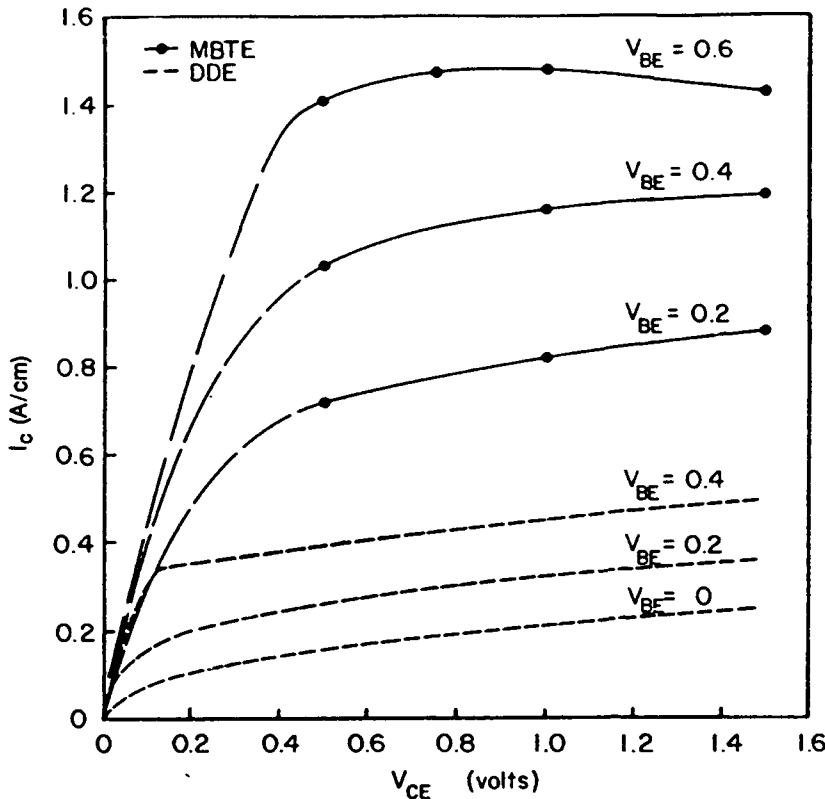
Figure 1.13 Collector current versus collector-emitter voltage for different values of base-emitter voltage [20].

A comparison between the total carrier density distribution along the center of the channel for drift and diffusion and MBTE solutions is shown in Fig. 1.14 [20] for $V_{CE} = 1.0$ V and $V_{BE} = 0.4$ V. The moment equation prediction for the Γ-valley carrier density is also shown.

As seen in Fig. 1.14 for the DDE simulations, the carrier density reaches a maximum between base contacts and displays a significant dipole layer. Here, with the velocity in saturation and the cross-sectional area at a minimum, the carrier density must increase to maintain current continuity. In the moment equation simulation, the constraints of current continuity are more complex. First, a decrease in the cross-sectional areas is, as in the DDE, accompanied by an increase in field along the channel. The field increase under both equilibrium and nonequilibrium conditions is qualitatively similar, as may be observed from Fig. 1.15 [20], which shows the potential distribution along the center of the PBT channel. However, consequent changes in electron temperature, both increasing and decreasing, lag behind the equilibrium state. This leads to velocity overshoot (Fig. 1.16 [20]) and a delay in electron transfer. As a result, for nearly the first half of the device, transport is almost exclusively Γ-valley transport. The implication is that if the Γ-valley carrier velocity increases with increasing field, then the product of density and cross-sectional area normal to current
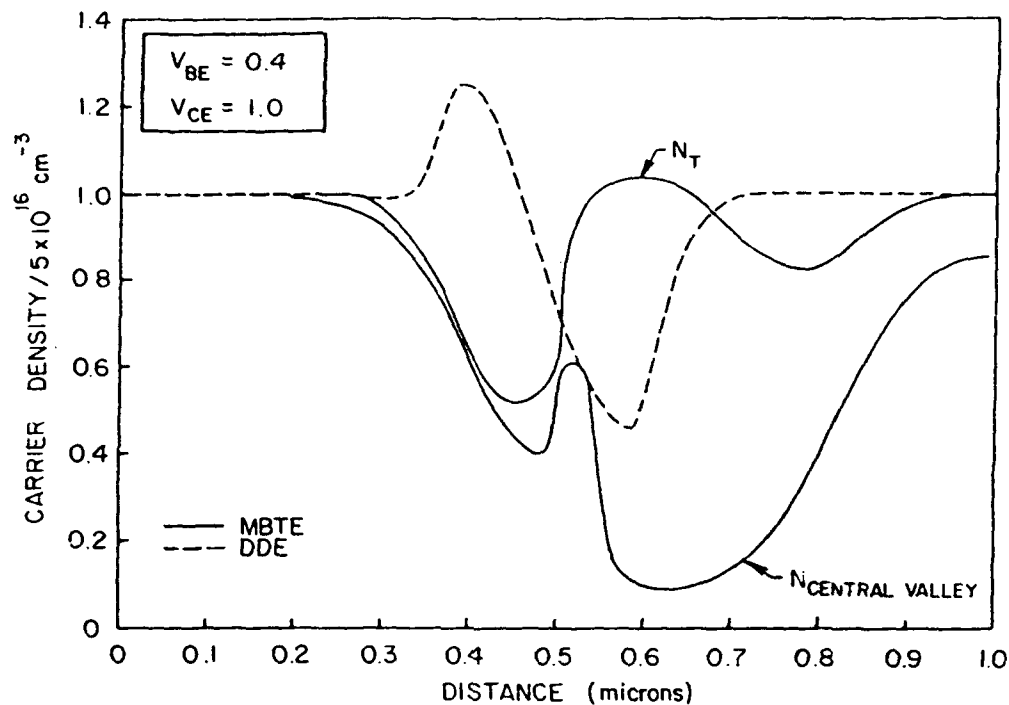
Figure 1.14 Carrier density versus distance along center of channel for the PBT [20].

flow must decrease to maintain current continuity. Since the velocity increases faster than the area decreases, the carrier density decreases.

At moderate bias levels, typical FET calculations show a decreasing field as the gate region is passed. This also occurs in the PBT. Now, as the cross-sectional area increases, the Γ-valley carriers exhibit a decrease in velocity. It must be noted, however, that for the parameters of the calcula-
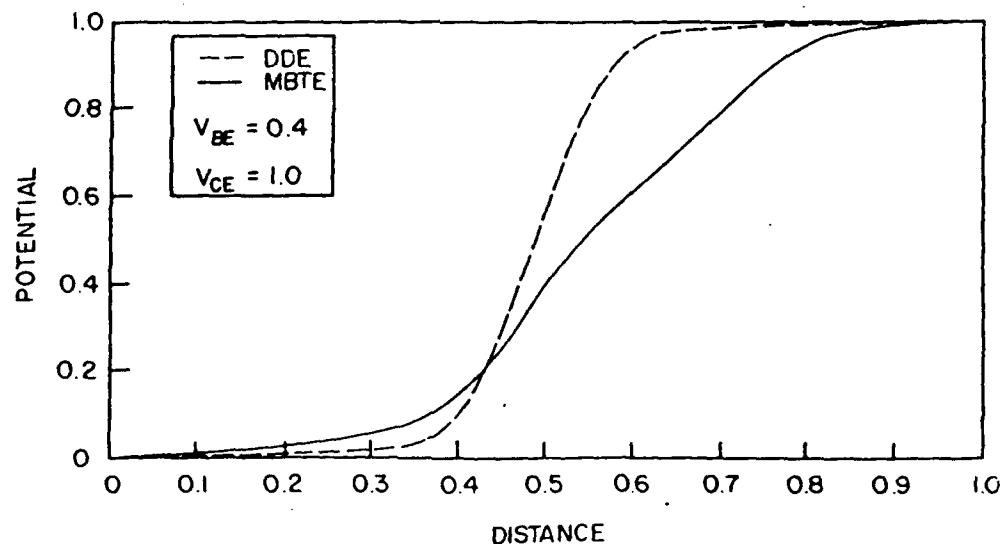

Figure 1.15 Potential versus distance along center of channel for the PBT [20].
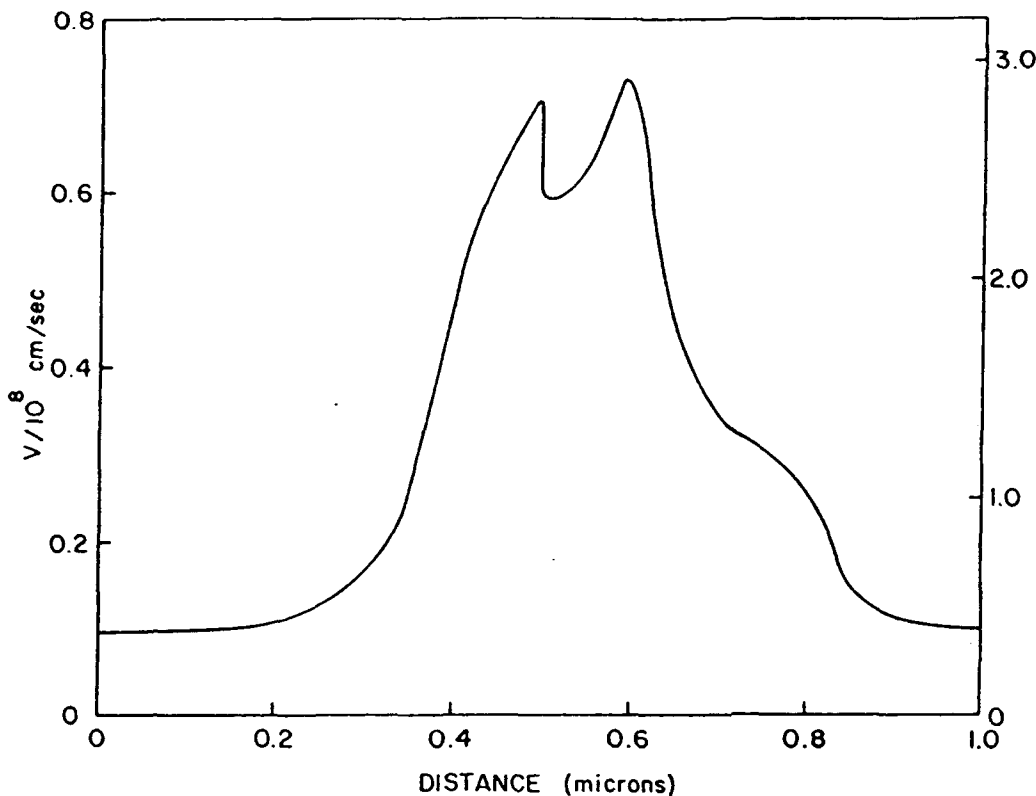
Figure 1.16 Γ-valley velocity along the center of the PBT channel [20].

tions the L-valley carriers make a negligible contribution to current. Thus, a decrease in carrier velocity results in a net increase in carrier concentration. However, initially, the decrease in field is not accompanied by a corresponding temperature decrease (as experienced in the uniform field calculations). Thus, the high Γ-valley temperature results in transfer to the L-valley, giving rise to the second minimum in the Γ-valley carrier density shown in Fig. 1.14. Further toward the drain, the field decreases. However, relaxation is incomplete and the field at the collector is not equal to the field at the emitter. Also note that the moment equation potential distribution gives rise to a slightly higher field upstream of the base, and a lower field, over a longer distance, downstream of the base compared with the drift and diffusion result. More significantly, the electron temperature at the collector exceeds that at the emitter. It is noted that as the field relaxes, the electrons transfer back to the central valley.

1.5 NONEQUILIBRIUM ELECTRON-HOLE TRANSPORT

Additional nonequilibrium studies were mentioned at the beginning of this chapter. This concerns nonequilibrium electron-hole transport, which for specificity was discussed for InGaAs. The details are considered below.

In this recent study, ensemble Monte Carlo studies were performed in which electrons were injected into p-type InGaAs [8]. In one case, the acceptor doping was 10^{17} cm^{-3}, and in the second case 5×10^{17} cm^{-3}. The calculations were at 300 K and the ratio of the injected electrons to the majority holes was taken to be 1:5. (Note: The ensemble Monte Carlo avoids any assumptions on the magnitudes of the energy and momentum exchange in an electron-hole [e-h] scattering process and the evolution of the electron and hole distribution functions.) The electrons and holes were assumed to be in equilibrium with the lattice when the electric field was switched on, and the band model consisted of three nonparabolic valleys for the conduction band and two parabolic light- and heavy-hole bands. The role of the light holes was suppressed in this study. The model includes the elastic acoustic phonon, impurity scattering using Ridley's model, alloy scattering, deformation potential, and intervalley and intravalley phonon scattering process. The e-h and screened carrier-phonon scattering are calculated from the expressions given in Osman and Ferry [22], using a self-consistent screening model. There is only one LO phonon mode in this
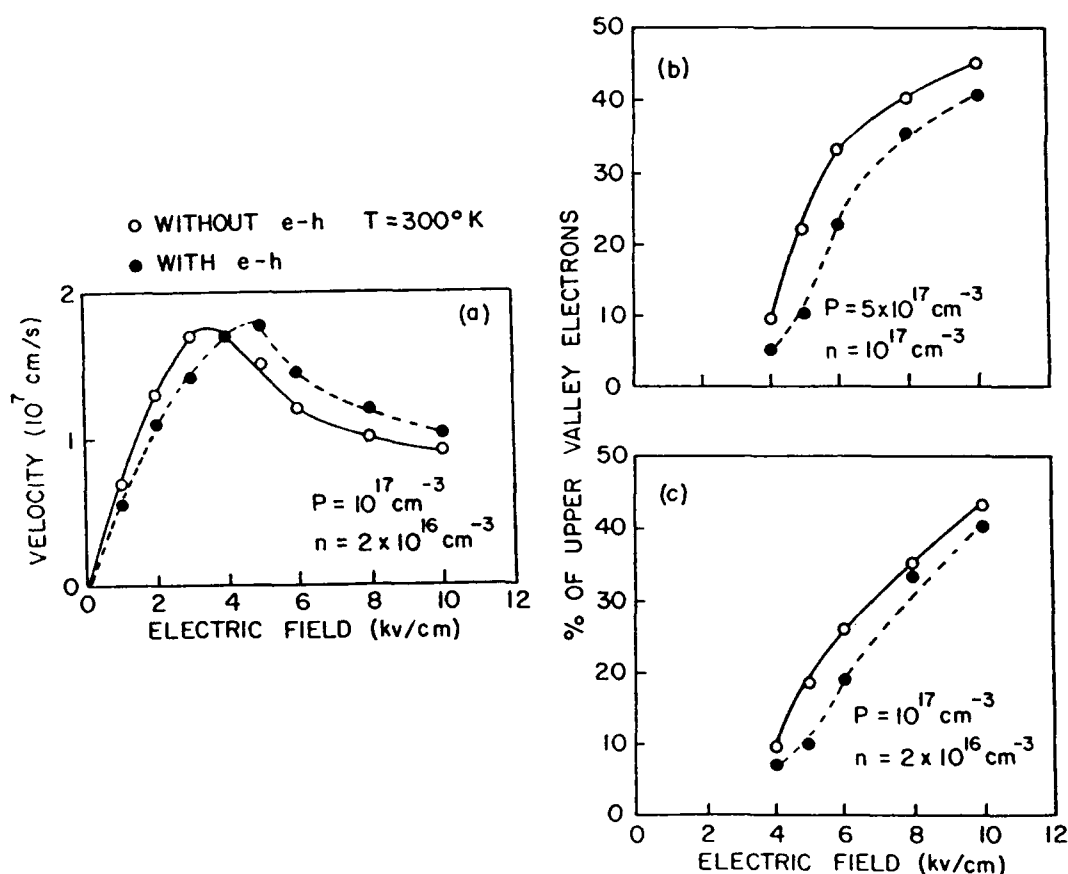

Figure 1.17   (a) Electron drift velocity in p-type InGaAs, (b–c) percentage of upper-valley electrons in p-type InGaAs [8] (Reprinted, with permission, from *Applied Physics Letters*.)

calculation. The electron transport parameters for $In_{0.53}Ga_{0.47}$ As are the same as those reported in Ahmed et al. [23]. However, for hole effective masses and deformation potential constants, appropriate to GaAs were used. The interaction between the L-valley electrons and the heavy holes was ignored.

The drift velocities of the electrons injected into the p-type InGaAs as a function of the applied electric fields are plotted in Figs. 1.4 and 1.17a [8], for doping levels of 5×10^{17} and 10^{17} cm^{-3}, respectively. The curves connecting the open circles in these figures correspond to situations in which the interaction with the mobile holes is ignored and only the interaction with the ionized acceptor impurities is taken into account.

When the interaction between the minority electrons and the mobile majority holes is taken into account, the electron velocities are lower for fields below 4 kV/cm compared with majority electrons, as can be seen from these figures. At these low fields, the electron transfer to the upper valleys is negligible, as shown in Fig. 1.17b [8], and the energy loss through e-h interaction is not significant because the rate at which the electrons gain energy from the electric field is small, as can be seen in Fig. 1.18 [8]. However, the e-h scattering, which has the same angular distribution as the impurity scattering, has the same effect on the electron mobility as doubling the doping level. Consequently, the mobility of the electrons is reduced, leading to lower velocities. As the electric field is increased, the fraction of electrons with enough energy to transfer to the upper valleys increases.
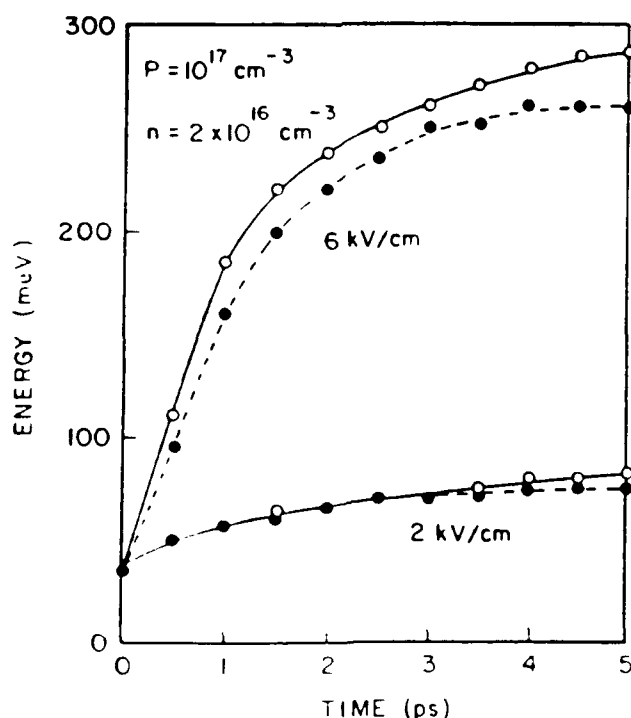

Figure 1.18   Electron energy in p-type InGaAs as a function of field [8].

THIS PAGE IS MISSING IN ORIGINAL DOCUMENT

22

For instance, from the equation of motion of the density matrix, for a system including mobile carriers and scattering centers, the first three moment equations have the following form [27]:

$$\langle\!\langle \dot{P}^{(0)} \rangle\!\rangle + \frac{1}{m}\frac{\partial}{\partial x}\langle\!\langle P^{(1)} \rangle\!\rangle = \frac{i}{\hbar}\langle\!\langle [H_s, P^{(0)}] \rangle\!\rangle \tag{12}$$

$$\langle\!\langle \dot{P}^{(1)} \rangle\!\rangle + \frac{1}{m}\frac{\partial}{\partial x}\langle\!\langle P^{(2)} \rangle\!\rangle = -\left(\frac{\partial V}{\partial x}\right)\langle\!\langle P^{(0)} \rangle\!\rangle + \frac{i}{\hbar}\langle\!\langle [H_s, P^{(0)}] \rangle\!\rangle \tag{13}$$

$$\langle\!\langle \dot{P}^{(2)} \rangle\!\rangle + \frac{1}{m}\frac{\partial}{\partial x}\langle\!\langle P^{(3)} \rangle\!\rangle = -2\left(\frac{\partial V}{\partial x}\right)\langle\!\langle P^{(1)} \rangle\!\rangle + \frac{i}{\hbar}\langle\!\langle [H_s, P^{(2)}] \rangle\!\rangle \tag{14}$$

where the $\langle\!\langle \ \rangle\!\rangle$ denote the quantum ensemble averages, and using Dirac notation, the operators of interest are of the form

$$P^{(0)} = |x_0\rangle\langle x_0| \tag{15}$$

$$P^{(1)} = (\tfrac{1}{2})(P|x_0\rangle\langle x_0| + |x_0\rangle\langle x_0|P) \tag{16}$$

$$P^{(2)} = (\tfrac{1}{2})^2(P|x_0\rangle\langle x_0| + 2P|x_0\rangle\langle x_0|P + |x_0\rangle\langle x_0|P^2) \tag{17}$$

$$P^{(3)} = (\tfrac{1}{2})^3(P^3|x_0\rangle\langle x_0| + 3P|x_0\rangle\langle x_0|P^2 + 3P^2|x_0\rangle\langle x_0|P$$
$$+ |x_0\rangle\langle x_0|P^3) \tag{18}$$

where P is the momentum operator. Note that the terms involving H_s incorporate dissipation. In a diagonal representation, the ensemble average of the first three operators breaks down into the following form:

$$\langle\!\langle P^{(0)} \rangle\!\rangle = \sum \rho_{ii} n_i(x_0) = n(x_0) \tag{19}$$

$$\langle\!\langle P^{(1)} \rangle\!\rangle = \sum \rho_{ii} n_i m v_i \equiv n(x_0) m v_d \tag{20}$$

$$\langle\!\langle P^{(2)} \rangle\!\rangle = \sum \rho_{ii} n_i m^2 [(v_i - v_d) + v_d]^2 - \frac{\hbar^2}{4}\sum \rho_{ii} n_i \frac{\partial^2}{\partial x^2}\ln n_i$$
$$= \Omega_{xx} + m^2 n v_d^2 - \frac{\hbar^2}{4}\sum \rho_{ii} n_i \frac{\partial^2}{\partial x^2}\ln n_i \tag{21}$$

where ρ_{ij} is the diagonal element of the density matrix, and

$$\Omega_{xx} = \sum \rho_{ii} n_i m(v_i - v_d)^2 \tag{22}$$

It is clear that with the exception of the third operator, which contains a term involving Planck's constant, the equations are of a classical form. Thus, in the simplest approximation, there appears to be a close similarity between

the classical moment equations and that obtained quantum mechanically. The difficulty is, of course, in solving these equations.

There is, however, an interesting situation to consider: that in which $\rho_{ii} = 1/N$ the system. In this case, the first two moment equations, including dissipation in momentum, reduce to

$$\frac{\partial n}{\partial t} + div(nv) = 0 \tag{23}$$

$$\frac{\partial nv}{\partial t} + \frac{\partial}{\partial x} nv_d^2 + \frac{nv}{\tau} = -\left(\frac{\partial V}{\partial x} + \frac{\partial Q}{\partial x}\right)\frac{n}{m} \tag{24}$$

where

$$Q = -\frac{\hbar^2}{2m} \frac{1}{\sqrt{n}} \frac{\partial^2 \sqrt{n}}{\partial x^2} \tag{25}$$

The quantity V represents an imposed barrier and the self-consistent energy associated with Poisson's equation. The potential Q [24] is density-dependent and tends to become significant near strong barriers, where the curvature of $\sqrt{n}$ will either enhance or diminish the imposed barrier. Tunneling and resonance arise from Q.

In multiple dimensions, these equations are subject to the constraint

$$\oint mv \cdot dx = nh \tag{26}$$

or in a gauge that includes a vector potential, the constraint

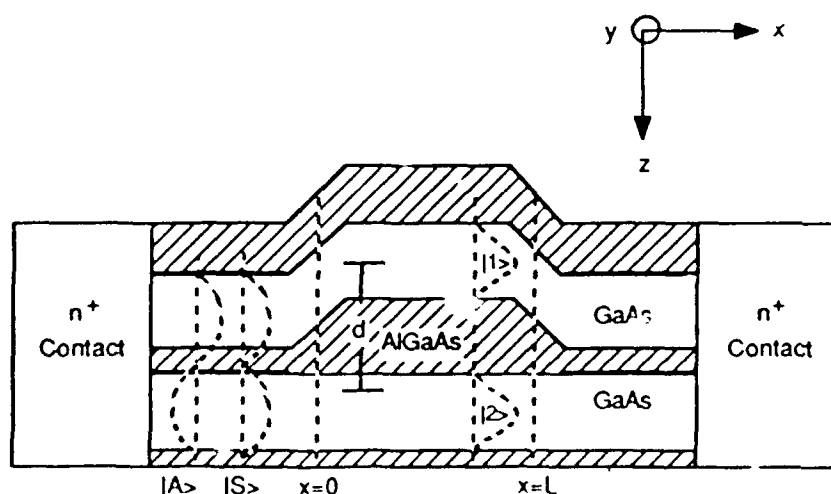$$\oint \left(mv + \frac{e}{c} \mathbf{A}\right) \cdot dx = nh \tag{27}$$

Figure 1.20 Configuration suitable for the Aharonov–Bohm constraint [29]. (Reprinted, with permission, from *Applied Physics Letters.*)

Presently, device systems are being constructed which are influenced by the constraint of Eq. (27), often called the Aharonov–Bohm condition [28]. In particular, structures are being constructed in which the path of an incident beam of electrons is split and then recombined. The split path lengths of the original beam are different, and under coherent reconstruction in which Eq. (27) is satisfied, conduction oscillations are anticipated. A structure originally proposed to deal with this is displayed in Fig. 1.20 [29].

ACKNOWLEDGMENTS

REFERENCES

1. J. B. Gunn, *Solid State Commun.* **1**, 88 (1963).

2. J. G. Ruch. *IEEE Trans. Electron Devices* **ED-19**, 652 (1972).

3. A. A. Ketterson, W. T. Masselink, J. G. Gedymin, J. Klem, C.-K. Peng, W. F. Kopp, H. Morkoç, and K. R. Gleason, *IEEE Trans. Electron Devices* **ED-33**, 564 (1986).

4. J. R. Chelikowsky and M. L. Cohen, *Phys. Rev. B: Solid State* [14] 556 (1976).

5. H. L. Grubin, *Lecture Notes in* "The Physics of Sub-Micron Semiconductor Devices," (H. L. Grubin, D. K. Ferry, and C. Jacoboni, eds.). Plenum Press, 1988.

6. V. L. Dalal, A. B. Dreeben, and A. Triano, *J. Appl. Phys.* **42**, 2864 (1971).

7. R. J. Degani, R. F. Leheny, R. E. Nahory, and J. P. Heritage, *Appl. Phys. Lett.* **39**, 569 (1981).

8. M. A. Osman and H. L. Grubin, *Appl. Phys. Lett.* **51**, 1812 (1987)

9. O. Madelung and M. Schulz, eds., "Numerical Data and Functional Relationships in Science and Technology/Landolt-Bornstein," Vo.. 22. Springer-Verlag, Berlin, 1987.

10. B. K. Ridley, "Quantum Processes in Semiconductors." Oxford Univ. Press (Clarendon), London and New York, 1982.

11. H. L. Grubin and J. P. Kreskovsky, *Physica* **134B + C**, 67 (1985).

12. H. L. Grubin, D. K. Ferry, and K. R. Gleason, *Solid-State Electron.* **23** 157 (1980).

13. P. N. Butcher, *Rep. Prog. Phys.* **30**, 97 (1967).

14. K. Yamaguchi, S. Asai, and H. Kodera, *IEEE Trans Electron Devices* **ED-23**, 1283 (1976).

15. M. Meyyappan, J. P. Kreskovsky, and H. L. Grubin, to be published.

16. M. P. Shaw, P. R. Solomon, and H. L. Grubin, *IBM J. Res. Dev.* **13**, 587 (1969).

17. E. M. Azoff, *Solid State Electron.* **30**, 913 (1987).
18. R. K. Froelich, *Avion. Lab. Rep.* **AFWAL-TR-82-1107,** 1982.
19. C. O. Bozler and G. D. Alley, *IEEE Trans. Electron Devices* **ED-27,** 6 (1980).
20. J. P. Kreskovsky, M. Meyyappan, and H. L. Grubin, *in* "Proceeding of NUMOS I" (J. J. H. Miller, ed.). Boole Press, 1987.
21. R. F. Cook and J. Frey, *COMPEL* **1,** 2 (1982).
22. M. A. Osman and D. K. Ferry, *J. Appl. Phys.* **61,** 5330 (1987).
23. S. R. Ahmed, B. R. Nag, and M. D. Roy, *Solid-State Electron.* **28,** 1193 (1985).
24. C. Philippidis, D. Bohm, and R. D. Kaye, *Nuovo Cimento Soc. Ital. Fis. B* [11] **71B,** 75 (1982).
25. G. J. Iafrate, H. L. Grubin, and D. K. Ferry, *J. Phys. (Orsay, Fr.)* **42,** C7-307 (1981).
26. E. Wigner, *Phys. Rev.* **40,** 749 (1932).
27. H. L. Grubin, to be published.
28. Y. Aharonov and D. Bohm, *Phys. Rev.* **115,** 485 (1959).
29. S. Datta, M. R. Mellcoh, S. Bandyopadyay, and M. S. Lundstrom, *Appl. Phys. Lett.* **48,** 487 (1986).

THE ROLE OF LONGITUDINAL-OPTICAL PHONONS IN NANOSCALE STRUCTURES

Michael A. Stroscio, Gerald J. Iafrate
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, North Carolina

K. W. Kim, M. A. Littlejohn
Department of Electrical and Computer
Engineering
North Carolina State University
Raleigh, North Carolina 27695

Harold L. Grubin
Scientific Research Associates, Inc.
Glastonbury, Connecticut 06033

V. V. Mitin, R. Mickevicius
Department of Electrical and Computer
Engineering
Wayne State University
Detroit, Michigan 48202

As device dimensions in nanoscale structures and mesoscopic systems are reduced, the characteristics and interactions of dimensionally-confined longitudinal-optical (LO) phonons deviate substantially from those of bulk polar semiconductors. This account emphasizes LO-phonon effects arising in three separate systems: Short-period AlAs-GaAs superlattices; rectangular GaAs quantum wires embedded in AlAs; and metal-semiconductor interfaces such as those in de Broglie wave quantum-interference devices.

I. INTRODUCTION

In nanoscale and mesoscopic systems, the effects of confinement on carriers have been studied extensively. However, to properly model carrier energy loss in nanoscale and mesoscopic systems, it is essential that calculations of carrier scattering by longitudinal-optical phonons take into account the fact that confinement also changes the strength and spatial properties of longitudinal-

optical phonons. As a first example of phonon-confinement effects, we shall emphasize the dominant role of interface-optical phonons on electron scattering in heterostructures with confinement lengths less than about 40 Angstroms.

Electron interactions with longitudinal-optical (LO) phonon modes in heterostructures are strongly affected by the changes in the Frohlich Hamiltonian caused by phonon confinement and localization, as well as by the changes in the electronic wave function due to the confining potential. The presence of heterointerfaces produces large changes in the dielectric constant and gives rise to the confinement of optical phonons in each layer (i.e., confines modes) and the localization in the vicinity of interfaces (i.e., interface modes or surface-optical (SO) modes). These confined and interface modes arise because the abrupt changes in dielectric constant near heterointerfaces make it impossible to satisfy the normal bulk dispersion relation which normally restricts LO phonon frequencies to those where the dielectric constant vanishes. There have been suggestions that interactions by the interface modes can be significant and that the scattering rate due to the confined modes can be considerably reduced in some structures compared to the bulk LO-phonon scattering rate.[1-3] Therefore, an appropriate treatment of the optical-phonon modes in quantized systems is essential for understanding electron transport in heterostructures. Recently, both macroscopic and microscopic approaches to electron-optical-phonon interactions in heterostructures [1-11] have been applied in theoretical treatments. Enhanced electron-SO-phonon (i.e., electron-interface-phonon) scattering in polar semiconductors with confining dimensions less than about 50 Angstroms has been indicated recently [12,13] on the basis of scalar-potential modes. Physically, this enhanced scattering rate is due to the interface phonon potential which increases approximately exponentially near heterointerfaces, where there is an abrupt change in the dielectric constant. The appropriateness of using a scalar potential to model the electron-SO-phonon interaction has been questioned.[14] Herein, we model the potential to calculate the ratio of electron-SO-phonon scattering to electron-confined-LO-phonon scattering in a short-period GaAs/AlAs superlattices.[15] The results provide a criterion which, in conjunction with experiments, can be used to examine the validity of the scalar-potential model.

II. MODEL AND PROCEDURE

Figure 1 shows a schematic drawing of the GaAs/AlAs short-period superlattice considered in this study. Each GaAs layer has width, a, and each AlAs layer has width, b. We will consider the case where a uniform electron current, J, is incident parallel to the superlattice heterojunctions. The ratio of electron-SO-phonon scattering to electron-LO-phonon scattering in the superlattice of Fig. 1 is calculated through the Fermi golden rule by treating the electron-Lo-phonon and electron-SO-phonon interaction Hamiltonians [4,6] as perturbation
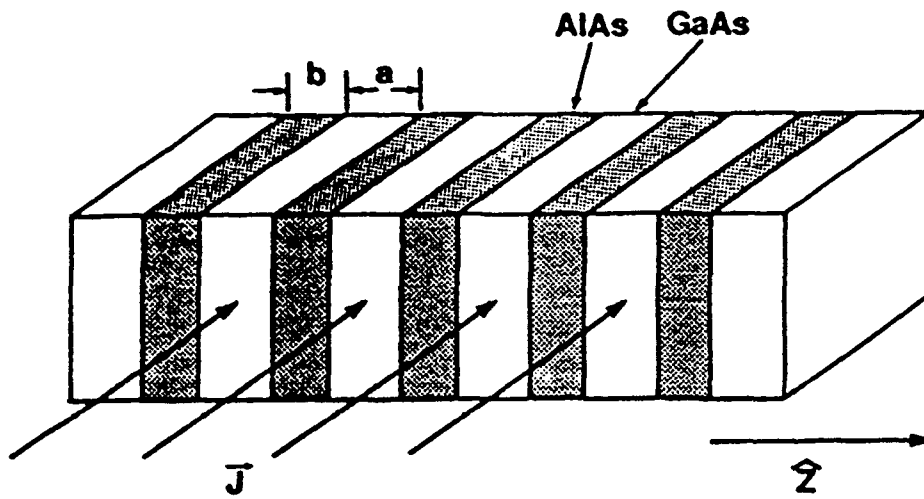
FIG 1.   Short-period GaAs/AlAs superlattice with a uniform electron curent parallel to the heterojunction interfaces.

Hamiltonians and by taking the electronic wavefunction in the z-direction as that given by Cho and Prucnal for the maximum edge of the first subband.[16]

To obtain an order-of-magnitude estimate of the ratio of electron-SO-phonon scattering to electron-LO-phonon scattering in the GaAs/AlAs superlattice of Fig. 1, we take the ratio of the Fermi golden rule transition probability for electron scattering from the symmetric AlAs-like interface mode to the corresponding transition rate for electron scattering from the lowest order confined phonon mode.

In Fig. 2, this ratio is plotted as a function of the parallel phonon wavevector times the quantum well width, qa, for a GaAs/AlAs short-period superlattice with a=b. For qa < π, the ratio is larger than unity. A principal factor contributing to the rapid increase for small values of aq (=X) is the exponential increase in the strength of the SO-phonon mode near the heterojunction. In agreement with our results, a surge in the strength of the S+ mode with a decreasing well-width, a, is observed in a recent low-temperature experiment,[17] which clearly demonstrates the dominance of the SO-phonon modes over the LO-phonon modes. These results are based on the implicit assumption that the uniform current, J, is used to prepare the electronic state given by Ref. 16. This assumption is reasonable if the electrons have incident energies significantly less than the GaAs/AlAs conduction band discontinuity and if the superlattice barrier widths are roughly 50 Angstroms or less. Furthermore, the frequency of the symmetric AlAs-like mode for a GaAs/AlAs heterojunction varies from about 45
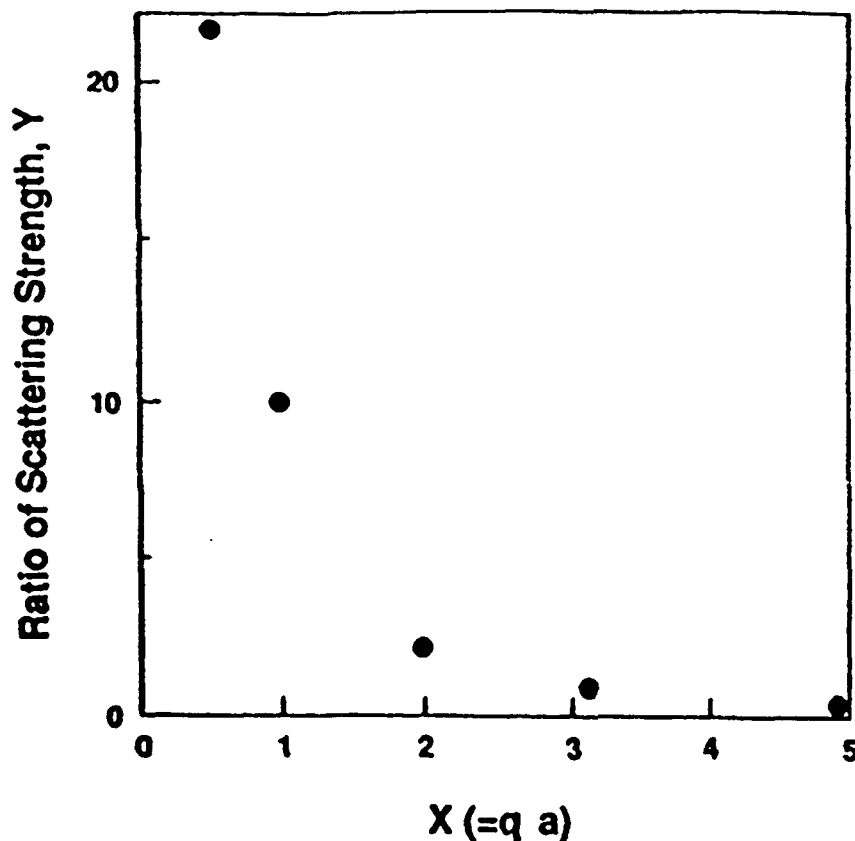
FIG. 2. Approximate ratio of electron-SO-phonon (S+ mode) scattering to electron-LO-phonon (m=1 confined mode) scattering.

meV for $qa > 6$ to 50 meV near $qa=0$ and the frequency of the lowest-order confined mode is approximately 36.2 meV. Thus, the scaling illustrated in Fig. 2 is expected to hold for a < 50 Angstroms and for incident electron energies in the range from about 50 meV to very roughly 150 meV. In this energy range, the phonon momentum, q, will be approximately equal to the incident electron momentum. Thus, we conclude that enhanced SO-phonon scattering will oc r in the regime of low-voltage (also, low-temperature in general) operation characteristic of mesoscale devices currently being considered by the electron device community.[18-20]

As a second example, of the nature of Lo-phonon confinement effects in polar semiconductor nanostructures, we consider Golden rule scattering rates for emission and absorption in rectangular GaAs quantum wires embedded in AlAs. In these calculations, the carriers are taken to be free in the x-direction and bound in the ground states of finitely deep quantum wells in the y- and z-directions. For a 40-Angstrom-by-40-Angstrom GaAs quantum wire of infinite length in the x-direction, we find strong one-dimensional density-of-states effects

in the rates for surface-optical (SO) phonon emission. The emission peaks for the SO-phonons, also known as interface-optical phonons, occur near the AlAs-like and GaAs-like LO-phonon frequencies. The confined and interface phonon modes used in obtaining these results are those considered previously by the authors.[6,9,12,21] Fig. 3 depicts the Golden rule scattering rates for the quantum wire. The strong SO-phonon emission rates near the AlAs-like and GaAs-like LO-phonon frequencies are more than an order of magnitude larger than the corresponding bulk LO phonon emission rates.[9,21] However, if the lateral dimensions of the quantum wire are varied sinusoidally in amplitude by only ten percent, the narrow one-dimensional density-of-states emission peaks have averaged values close to those of bulk carrier—LO-phonon interaction rates.[21] In this case, we learn that carrier—LO-phonon scattering rates may deviate
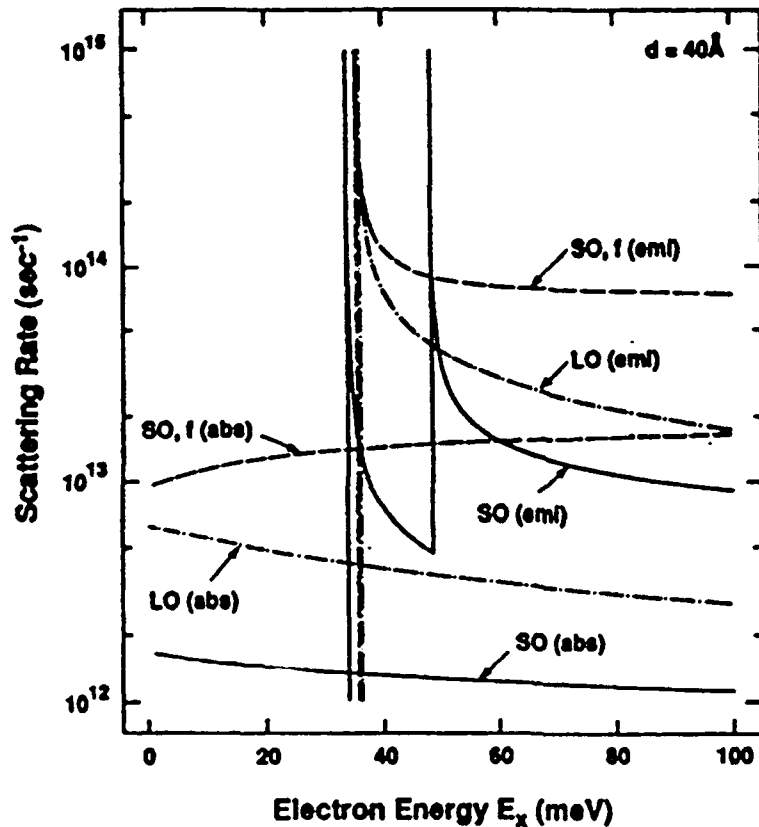
FIG.3. Scattering rates at 300 K as a function of electron energy in the longitudinal direction. The size of quantum wire is 40 Angstroms by 40 Angstroms. The SO-phonon scattering rate is denoted as SO, f for the free-standing quantum wire. and as SO for the GaAs/AlAs quantum wire, respectively, while LO represents the rate of electron scattering with confined LO-phonon modes. Phonon emission (absorption) process is represented as emi (abs), respectively.

substantially from the corresponding wires; however, in the more realistic case of variable quantum wire dimensions, one-dimensional effects are suppressed.

As a third and final example of how optical phonon properties in nanoscale structures may deviate substantially from those of bulk optical phonons, we consider a metal-semiconductor interface for the case where the semiconductor thickness is restricted to nanoscale dimensions. In this analysis, we find that the nature of the interface phonon modes at the metal-semiconductor interface are modified substantially for two reasons: (1) at the metal-semiconductor interface, only those phonon modes with Frohlich potentials which approach zero at the interface are allowed; and (2) the magnitude of the Frohlich potential for the surviving interface phonon mode decreases exponentially with the product of the phonon wavevector and thickness of the semiconductor slab. As demonstrated previously,[4,6] the components of the optical phonon polarization vectors for dimensionally-confined modes have opposite parities for polarization components normal to the semiconductor interface and for polarization components parallel to the interface. Since the average tangential components of the optical-phonon electric field must vanish at the metal interface and since the phonon polarization vector is proportional to the phonon electric field, it follows that only optical phonon modes with even polarization vectors normal to the metal-semiconductor interface will survive. Furthermore, since the laplacian of the Frohlich potential for a given phonon mode is proportional to the divergence of the polarization vector for that mode, it follows that only optical phonon modes with odd Frohlich potentials satisfy the metal-semiconductor interface. In particular, of the well-known symmetric and antisymmetric interface modes,[21] only the antisymmetric mode will survive. In the limit of no dispersion, the antisymmetric interface mode has a Frohlich potential which scales as one divided by the square root of $\cosh(qd) \bullet \sinh(qd)$ where q is the magnitude of the phonon wavevector parallel to the metal-semiconductor interface and d is the thickness of the semiconductor slab. For small values of qd, the Frohlich potential of the antisymmetric interface mode then decreases exponentially with qd. Hence, we conclude that the carrier-optical-phonon scattering rate decreases to zero at the metal-semiconductor interface and that the strength of the Frohlich potential for the antisymmetric interface mode in the semiconductor slab decreases exponentially with qd. (It is important to note that these results depend upon having a perfect metal-semiconductor interface since even a thin oxide at the metal-semiconductor interface is sufficient to isolate the semiconductor optical-phonon modes from the metal.) The impact of these observations on the physics of nanoscale structures such as the de Broglie wave interference device is clear: carrier-optical-phonon scattering will be reduced near metal-semiconductor interfaces. Additionally, Schottky barrier heights at metal-semiconductor interfaces in nanoscale structures may deviate from those of metal-semiconductor structures with semi-infinite metal and semiconductor

384

regions; the extent of this deviation will depend on the relative importance of phonon scattering and other factors such as interface states which determine the equilibrium Schottky barrier height.

III. CONCLUSION

In conclusion, the characteristics and interactions of dimensionally-confined optical phonon modes deviate substantially from those of bulk polar semiconductors. In particular, theory and experiment point to greatly enhanced carrier—LO-phonon scattering rates when dimensional constraints optimize interface mode strengths. For the case of enhanced carrier—LO-phonon emission rates in rectangular quantum wires, our analysis indicates the one-dimensional density-of-states effects may be greatly reduced as a result of minor fluctuations in the laterial dimensions of the quantum wire. Finally, we conclude that necessary restrictions in phonon symmetries at metal-semiconductor interfaces will modify carrier-LO-phonon scattering rates in nanostructures where charge transport occurs in thin polar semiconductor regions in intimate contact with metals.

REFERENCES

1. R. Lassnig, Phys. Rev. B **30**, 7132 (1984).
2. F. A. Riddoch and B. K. Ridley, Physica B+C **134B**, 342 (1985).
3. N. Sawaki, I. Phys. C: Solid State Phys. 19, 4965 (1986).
4. J. J. Licari and R. Evrard, Phys. Rev. B **15**, 2254 (1977).
5. L. Wendler, Phys. Status Solidi B **129**, 513 (1985).
6. N. Mori and T. Ando, Phys. Rev. B **40**, 6175 (1989).
7. R. Chen, D. L. Lin, and T. F. George, Phys. Rev. B **41**, 1435 (1990).
8. B. K. Ridley, Phys. Rev. B **39**, 5282 (1989).
9. M. A. Stroscio, Phys. Rev. B **40**, 6428 (1989).
10. K. Huang and B. Zhu, Phys. Rev. B **38**, 13377 (1988).
11. S. Rudin and T. L. Reinecke, Phys. Rev. B **41**, 7713 (1990).

12. M. S. Stroscio, K. W. Kim, M. A. Littlejohn, and H. Chuang, Phys. Rev. B 42, 1488 (1990); K. W. Kim, M. A . Stroscio, A. Bhatt, R. Mickevicius and V. Mitin, J. Appl. Phys. (to be published).

13. S. Teitsworth and P. Turley, private communication.

14. B. K. Ridley and M. Babiker, private communication.

15. M. A. Stroscio, Gerald J. Iafrate, K. W. Kim, M. A. Littlejohn, H. Goronkin, and G. N. Maracas, Appl. Phys. Lett., in press (1991).

16. H.-S. Cho and P. R. Prucnal, Phys. Rev. B **36**, 3237 (1987).

17. K. T . Tsen and H. Morkoc, Private communication.

18. K. Sols, M. Macucci, U. Ravaioli, and K. Hess, Appl. Phys. Lett. **54**, 350 (1989); F. Sols and M. Macucci, Phys. Rev. B **41** 11887 (1990).

19. H. Sakaki, Jpn. J. Appl. Phys. **28**, L314 (1989).

20. G. J. Iafrate, in *Gallium Arsenide Technology*, edited by D. K. Ferry (Sams, Indianapolis, 1985), Chapter 12.

21. K. W. Kim and M. A. Stroscio, J. Appl. Phys. **68**, 6289 (1990).

Dramatic reduction in the longitudinal-optical phonon emission rate in polar-semiconductor quantum wires

By MICHAEL A. STROSCIO†, K. W. KIM‡, GERALD J. IAFRATE†,
MITRA DUTTA§ and HAROLD L. GRUBIN‖

† U.S. Army Research Office, P.O. Box 12211,
Research Triangle Park, North Carolina 27709, U.S.A.
‡ Department of Electrical and Computer Engineering,
North Carolina State University, Raleigh, North Carolina 27695, U.S.A.
§ U.S. Army Electronics Technology and Devices Laboratory,
Ft Monmouth, New Jersey 07703, U.S.A.
‖ Scientific Research Associates, Inc., 50 Nye Road,
Glastonbu y, Connecticut 06033, U.S.A.

ABSTRACT

Novel quantum-effect polar-semiconductor structures underlie technologies portending dramatic enhancements in the capability to process information orders of magnitude faster than is possible currently. In many embodiments of these quantum-effect structures, charges are transported in quasi-one-dimensional quantum wires which must support the transport of charges at high mobilities. However, it has recently been demonstrated that the longitudinal-optical (LO) phonons established at quantum-wire interfaces lead to dramatic enhancements in carrier–phonon interactions and concomitant degradation in carrier mobility. This letter demonstrates that phonon modes may be tailored through the judicious use of metal–semiconductor interfaces in such a way as to dramatically reduce unwanted emission of interface LO phonons and, consequently, to lead to the achievement of high quantum-wire mobility.

Modern fabrication techniques for making semiconductor structures with nanometre-scale dimensional features have been essential in leading to fundamental discoveries (Esaki 1974, Stein, von Klitzing and Weimann 1983) as well as in stimulating concepts for future information-processing systems based on the exploitation of quantum effects occurring in such structures (Capasso and Datta 1990). As originally proposed by Sakaki (1980), the predicted high mobilities of quasi-one-dimensional wire-like regions of a semiconducting material underlie many proposed quantum-wire system concepts (Luryi and Capasso 1985, Sakaki 1989) and have resulted in pioneering efforts leading to the actual fabrication of quantum wires (Watt, Sotomayor Torres, Arnot and Beaumont 1990). Furthermore, quantum-wire structures have been essential to seminal studies underlying quantum-wire laser concepts (Tsuchiya et al. 1989) and quantum-coupled electronic systems (Reed et al. 1988). Recently, however, theoretical studies of the interaction between longitudinal-optical (LO) phonons and carriers in a polar-semiconductor quantum wire (Stroscio 1989) have revealed the presence of discrete LO phonon modes similar to those identified for polar-semiconductor quantum wells (Kliewer and Fuchs 1966, Licari and Evrard 1977, Mori and Ando 1989, Kim and Stroscio 1990). As for the case of quantum wells, interface LO phonons are established at the semiconductor–semiconductor boundaries of quantum wires (Kim et al. 1991). In addition, it has been shown that, for carrier

energies in excess of the interface LO-phonon energy, the inelastic scattering caused by carrier–interface-phonon interactions dominates over other scattering mechanisms when confinement occurs on a scale for about 40 Å or less (Stroscio *et al.* 1991). In this letter, it is demonstrated that establishing a metal–semiconductor interface at the lateral boundaries of polar-semiconductor quantum wires introduces a set of boundary conditions that dramatically reduces or eliminates unwanted carrier energy loss caused by the interactions with interface LO-phonon modes.

Motivated by the recently demonstrated technology for the epitaxial growth of metals in intimate contact with polar semiconductors (Harbison *et al.* 1988, 1989, Guivarc'h *et al.* 1989), the dielectric continuum model of interface-phonon modes has been applied to determine the carrier–interface-phonon interaction Hamiltonian near a semi-infinite metal–polar-semiconductor interface (Stroscio *et al.* 1992). Detailed microscopic calculations of interface modes in polar semiconductors (Rücher, Molinari and Lugli 1991) indicate that the dielectric continuum model provides an accurate formalism for modelling carrier–interface-phonon interactions. As summarized by Mori and Ando (1989), the components of the optical polarization vectors for confined and interface phonons have opposite parities for polarization components normal to the semiconductor interface and for polarization components parallel to the interface. Since the average tangential components of the optical-phonon electric field must vanish at the metal interface (it may survive at a semiconductor–semiconductor interface) and since the phonon polarization vector is proportional to the phonon electric field, it follows that only optical-phonon modes with even polarization vectors normal to the metal–semiconductor interface will survive. Furthermore, since the Laplacian of the Fröhlich potential for a given phonon mode is proportional to the divergence of the polarization vector for that mode, it follows that only optical-phonon modes with odd Fröhlich potentials satisfy the metal–semiconductor interface conditions. Hence, only those confined and interface modes of odd potential satisfy the correct boundary conditions at the metal–semiconductor interface. In particular, of the well known symmetric and antisymmetric interface modes, only the antisymmetric modes will survive; then straightforwardly from the formulation of Kim and Stroscio (1990), it follows that

$$H_A = \sum_q \left(\frac{\hbar \exp^2 L^{-2}}{\varepsilon_0(\partial/\partial\omega)[\kappa_1(\omega)\coth(qd/2) + \kappa_2(\omega)]} \right)^{1/2}$$

$$\times \frac{1}{(2q)^{1/2}} \exp(i\mathbf{q}\cdot\rho)(a_\mathbf{q} + a^\dagger_{-\mathbf{q}}) \frac{\sinh(qz)}{\sinh(qd/2)}, \quad 0 < z < \frac{d}{2},$$

where H_A denotes the antisymmetric electron–interface-phonon interaction Hamiltonian in a polar-semiconductor slab of dielectric function $\kappa_1(\omega)$, bounded for $z < 0$ with a semi-infinite metal and for $z > d/2$ with a semi-infinite region of dielectric function $\kappa_2(\omega)$. In this result, $\mathbf{q}$ is the component of the interface-phonon wave-vector parallel to the semiconductor heterojunctions, ρ is the pair of spatial coordinates parallel to the planes of the semiconductor heterojunctions, z is the coordinate orthogonal to ρ, L is the normalization length, $a_\mathbf{q}$ and $a^\dagger_{-\mathbf{q}}$ are the usual phonon annihilation and creation operators, and ε_0 is the permittivity of vacuum. In the dispersion regions where $\mathbf{q}$ does not depend strongly on ω, H_A scales with $qd/2$ as $[\cosh(qd/2)\sinh(qd/2)]^{-1/2}$ or equivalently as $2\exp(-qd/2)$; thus, in the dispersionless region (away from $q = 0$), the strength of the electron–interface-phonon interaction decreases exponentially as $d/2$ becomes large. The impact of these observations on the

physics of nanostructures is clear: carrier–optical-phonon scattering will be reduced near metal–semiconductor interfaces.

For rectangular polar-semiconductor quantum wires embedded in a metal, the only confined and interface-LO-phonon modes satisfying the metal–semiconductor boundary conditions are the modes with odd potentials at all boundaries of the quantum wire. For the case of confined modes, the electron–LO-phonon Hamiltonian is precisely that given previously by Stroscio (1989) since all the modes derived earlier are odd at the quantum-wire interfaces; thus all results of Stroscio (1989) hold for the case of a rectangular polar-semiconductor quantum wire embedded in a metal. For the case of interface modes, there are no allowed solutions since it is impossible to satisfy the metal–semiconductor boundary conditions simultaneously at two of the quantum-wire metal–semiconductor interfaces. While the demonstration holds rigorously only for quantum wires with rectangular cross-sections, it is clear that the same conclusion will be substantially correct for a quantum wire of arbitrary cross-section since the interface modes with long-range Coulomb character must be reduced severely by invoking metal–semiconductor boundary conditions at two opposing locations on the quantum-wire cross-section.

The conclusion that carrier–interface-phonon interactions will be absent in quantum wires surrounded by metal is critical in maintaining high-mobility transport in quasi-one-dimensional quantum-wire structures. Indeed, the recent theoretical and experimental findings (Tsen *et al.* 1991) indicate that the strength of the carrier–interface-phonon scales approximately inversely with the thickness of the confining region. However, it is precisely the regime of the smallest dimensional scales which has portended the greatest advances in the novel application and discovery of quantum effects in semiconductor structures. As demonstrated in this letter, this highly unsatisfactory situation for polar-semiconductor quantum wires can be circumvented by eliminating interface-optical-phonon modes through the judicious use of metal–semiconductor interfaces.

References

Capasso, F., and Datta, S., 1990, *Phys. Today*, **74**, 75.

Esaki, L., 1974, *Rev. mod. Phys.*, **46**, 237.

Guivarc'h, A., Caulet, J., Guenais, B., Ballini, Y., Guerin, R., Poudoulec, A., and Regreny, A., 1989, *J. Cryst. Growth*, **95**, 427.

Harbison, J. P., Sands, T., Tabatabaie, N., Chan, W. K., Florez, L. T., and Keramidas, V. G., 1988, *Appl. Phys. Lett.*, **53**, 1717; 1989, *J. Cryst. Growth*, **95**, 425.

Kim, K. W., and Stroscio, M. A., 1990, *J. appl. Phys.*, **68**, 6289.

Kim, K. W., and Stroscio, M. A., Bhatt, A., Mickevicius, R., and Mitin, V. V., 1991, *J. appl. Phys.*, **70**, 319.

Kliewer, K. L., and Fuchs, R., 1966, *Phys. Rev.*, **150**, 573.

Licari, J. J., and Evrard, R., 1977, *Phys. Rev.* B, **15**, 2254.

Luryi, S., and Capasso, F., 1985, *Appl. Phys. Lett.*, **47**, 1347.

Mori, N., and Ando, T., 1989, *Phys. Rev.* B, **40**, 6175.

Reed, M. A., Randall, J. N., Aggrawal, R. J., Matyi, R. J., Moore, T. M., and Wetsel, A. E., 1988, *Phys. Rev. Lett.*, **60**, 535.

RÜCHER, H., MOLINARI, E., and LUGLI, P., 1991, *Phys. Rev.* B, **44**, 3463.

SAKAKI, H., 1980, *Jap. J. appl. Phys.*, **19**, L735; 1989, *Jap. J. appl. Phys.*, **28**, L314.

STEIN, D., VON KLITZING, K., and WEIMANN, G., 1983, *Phys. Rev. Lett.*, **51**, 130.

STROSCIO, M. A., 1989, *Phys. Rev.* B, **40**, 6428.

STROSCIO, M. A., IAFRATE, G. J., KIM, K. W., LITTLEJOHN, M. A., GORONKIN, H., and MARACAS, G. N., 1991, *Appl. Phys. Lett.*, **59**, 1093.

STROSCIO, M. A., IAFRATE, G. J., KIM, K. W., LITTLEJOHN, M. A., GRUBIN, H. L., MITIN, V. V., and MICKEVICIUS, R., 1992, *Nanostructure and Mesoscopic Systems*, edited by M. A. Reed and P. Kirk (Boston: Academic Press) (to be published).

TSEN, K. T., SMITH, D. S., TSEN, S.-C. Y., KUMAR, N. S., and MORKOC, H., 1991, *J. appl. Phys.*, **70**, 418.

TSUCHIYA, M., GAINES, J. M., YAN, R. H., SIMES, R. J., HOLTZ, P. O., COLDREN, L. A., and PETROFF, P. M., 1989, *Phys. Rev. Lett.*, **62**, 466.

WATT, M., SOTOMAYOR TORRES, C. M., ARNOT, H. E. G., and BEAUMONT, S. P., 1990, *Semicond. Sci. Technol.*, **5**, 285.

A NEW NONPARABOLIC HYDRODYNAMIC MODEL WITH QUANTUM CORRECTIONS

D. L. Woolard[a], M. A. Stroscio[b] M. A. Littlejohn[a,b],

R. J. Trew[a] and H. L. Grubin[c]

[a]Electrical and Computer Engineering Department
North Carolina State University
Raleigh, North Carolina 27695-7911

[b]U.S. Army Research Office
Research Triangle Park, North Carolina 27709

[c]Scientific Research Associates Inc.
Glastonbury, Connecticut, 06033-6058

Abstract - *This paper presents a new hydrodynamic transport model with nonparabolic conduction bands and quantum correction terms. For the first time solutions for the full quantum balance equations, applied to an ultrasmall electron device, are presented.*

1. Introduction

Requirements for faster electronics have produced smaller devices influenced by quantum interference effects which can not be modeled by classical theory. These devices operate under nonequilibrium conditions where the average carrier energy reaches many times its equilibrium value. To address these problems several approaches [1] have been considered. This paper describes a preliminary investigation into solutions for electron density, average velocity and average energy for an ultrasmall electronic structure under the conditions of nonclassical electron transport, using an approach previously described as quantum hydrodynamics [2].

2. The Nonparabolic Model

First, we present a brief summary of a unique form of the hydrodynamic transport model applicable to nonparabolic conduction bands. A complete description of this model will be formalized elsewhere [3]. Our hydrodynamic model was developed by studying moments of the Boltzmann transport equation [4]. This process was achieved using the moment operators $\Psi_1 = 1$, $\Psi_2 = u(k) = \frac{\hbar k}{m(k)}$, and $\Psi_3 = \frac{1}{2} m(k) u(k) \cdot u(k) \approx E_c(k)$. Here, the Kane dispersion relation, $\frac{\hbar^2 k^2}{2m^*} = E_c(1 + \alpha E_c)$, has been used to define $u(k)$ and $m(k) = m^* \sqrt{1 + \frac{2\alpha \hbar^2 k^2}{m^*}}$. These moment operators were chosen because they lead to a form which can be manipulated more easily and one in which simplifying approximations can be seen more clearly. Performing the moment process and simplying to first order yields immediately the

60

results:

$$\frac{\partial n}{\partial t} = -\nabla_r \cdot (n\mathbf{v}) \tag{1}$$

$$\frac{\partial \mathbf{v}}{\partial t} = -\mathbf{v} \cdot \nabla_r \mathbf{v} + \frac{\mathbf{F}}{m^{**}} - \frac{1}{nm^*}\nabla_r \cdot [\mathbf{P_v}] \tag{2}$$

$$\frac{\partial w}{\partial t} = -\mathbf{v} \cdot \nabla_r \ w + \mathbf{F} \cdot \mathbf{v} - \frac{1}{n}\nabla_r \cdot (\mathbf{v} \cdot [\mathbf{P_w}] + \mathbf{q}) \tag{3}$$

where m^{**}, $[\mathbf{P_v}]$, $[\mathbf{P_w}]$ and q have integral definitions [3] which depend on f. At this point additional assumptions or relations are necessary to close the mathematical system . A familiar approach is to use a displaced Maxwellian distribution function [7]. However, to resolve the dilemma of two distinctly different effective pressure tensors $\mathbf{P_v}$ and $\mathbf{P_w}$, (for parabolic bands and displaced Maxwellian, $[\mathbf{P_v}] = [\mathbf{P_w}] = nk_BT[\mathbf{I}]$), we use stationary Monte Carlo calculations and physical intuition to suggest the following distribution function

$$f = n[\frac{m_w(T_w)}{2\pi k_B T_w}]^{\frac{3}{2}} \exp[\frac{-m_w(T_w)|\mathbf{u} - \mathbf{v}|^2}{2k_B T_w}] \tag{4}$$

as a constitutive relation to close the moment equations. Here T_w has been chosen to replace T because $\frac{3}{2}k_B T_w$ approximates the effective thermal energy well for the stationary transport case, and $m_w(T_w)$ has been introduced because we expect a nonconstant effective mass strongly dependent on T_w. Using Equation (4) in the definitions for the transport parameters, and limiting the analysis to first and second orders in T_w, we arrive at the the supplemental relations; $w = \frac{3}{2}k_B T_w + \frac{m^{**}(T_w)}{2}\mathbf{v}\cdot\mathbf{v}$, $m^{**}(T_w) = m^*(1 + 3\alpha k_B T_w)$, $[\mathbf{P_v}] = \frac{m^*}{m^{**}}nk_B T_w[\mathbf{I}]$, $[\mathbf{P_w}] = nk_B T_w[\mathbf{I}]$ and $r_i = 5\alpha(\frac{m^*}{m^{**}})^2(k_B T_w)^2 nu_i$. These relations can be used to close Equations (1) - (3) expressing them in terms of electron density n, average velocity v and average energy w.

3. The Quantum Correction Terms

Since we desire a general model suitable for studying ultrasmall devices, an approach is illustrated to develop some quantum corrections for our semi-classical nonparabolic model equations. Grubin and Kveskovsky [2] have previously presented a detailed derivation which will achieve this goal. However, for our initial investigations we use a slightly more restricted version of their quantum hydrodynamic equations. Specifically, we neglect any spatial or nonparabolic effects on the electron effective mass and ignore any deficiency of the general quantum distribution function to agree with Fermi statistics (low temperature effects). We begin with the general moment equations of Stroscio [5]. The collisionless one-dimensional forms of the first three moments are:

$$\frac{\partial n}{\partial t} = -\frac{1}{m^*}\frac{\partial}{\partial x}(n\mathbf{p}_d) \tag{5}$$

$$\frac{\partial \mathbf{p}_d}{\partial t} = -\frac{\partial}{\partial x}(\frac{\mathbf{p}_d^2}{2m^*} + U_{eff}) - \frac{1}{nm^*}\frac{\partial}{\partial x}(n\langle(\mathbf{p} - \mathbf{p}_d)\rangle) \tag{6}$$

$$\frac{\partial}{\partial t}((\mathbf{p}-\mathbf{p}_d)^2)) = -\frac{1}{nm^*}\frac{\partial}{\partial x}(n\langle(\mathbf{p}-\mathbf{p}_d)^3\rangle) - \frac{\mathbf{p}_d}{nm^*}\frac{\partial}{\partial x}(n\langle(\mathbf{p}-\mathbf{p}_d)^2\rangle)$$

$$-\frac{2}{m^*}\frac{\partial}{\partial x}(n\langle(\mathbf{p}-\mathbf{p}_d)^2\rangle)\frac{\partial \mathbf{p}_d}{\partial x} + \frac{\mathbf{p}_d\langle(\mathbf{p}-\mathbf{p}_d)^2\rangle}{nm^*}\frac{\partial n}{\partial x} \qquad (7)$$

where n is the electron density, $\mathbf{p}$ is the single electron momentum, $\mathbf{p}_d$ is the classical average momentum and U_{eff} is the total effective electric potential. If we follow Ref. 2 and use the momentum displaced nonequilibrium Wigner distribution function,

$$f_W = \frac{n}{N}\exp\left[\frac{-\beta(\mathbf{p}-\mathbf{p}_d)^2}{2m^*}\right]\left(1+\frac{\gamma}{3}(1-\frac{\beta}{m^*}(\mathbf{p}-\mathbf{p}_d)^2 x)\frac{\partial(\frac{1}{n}\frac{\partial n}{\partial x})}{\partial x}\right) \qquad (8)$$

of Ancona and Iafrate [6], we obtain $n\langle(\mathbf{p}-\mathbf{p}_d)^2\rangle = \frac{m^*}{\beta}(1-\frac{2\alpha}{3\beta}\frac{\partial(\frac{1}{n}\frac{\partial n}{\partial x})}{\partial x})$ and $n\langle(\mathbf{p}-\mathbf{p}_d)^3\rangle = 0$ where $\alpha = \frac{\hbar^2\beta^2}{8m^*}$, $\beta = \frac{1}{k_BT}$ and $\gamma = \frac{\alpha}{\beta}$. These results can be used with Equations (5) - (7) to yield the collisionless quantum hydrodynamic equations:

$$\frac{\partial n}{\partial t} = -\frac{\partial(nv)}{\partial x} \qquad (9)$$

$$\frac{\partial v}{\partial t} = -v\frac{\partial v}{\partial x} - \frac{1}{m^*}\frac{\partial}{\partial x}(U_{eff}+\frac{Q}{3}) - \frac{1}{nm^*}\frac{\partial(nk_BT)}{\partial x} \qquad (10)$$

$$\frac{\partial w}{\partial t} = -v\frac{\partial w}{\partial x} - v\frac{\partial}{\partial x}(U_{eff}+\frac{Q}{3}) - \frac{1}{n}\frac{\partial(nvk_BT)}{\partial x} + \frac{\hbar^2}{12m^*}\frac{\partial}{\partial x}(\frac{1}{n}\frac{\partial n}{\partial x})\frac{\partial v}{\partial x} \qquad (11)$$

where $w = \frac{3}{2}k_BT + \frac{1}{2}m^*v^2 - \frac{\hbar^2}{24m^*}\frac{\partial}{\partial x}(\frac{1}{n}\frac{\partial n}{\partial x})$ and $Q = -\frac{\hbar^2}{2m^*}\frac{1}{\sqrt{n}}\frac{\partial^2\sqrt{n}}{\partial x^2}$. Comparing the above quantum equations to the semiclassical nonparabolic hydrodynamic Equations (1) - (3), approximating T_w by T and using the supplemental relations of Section 2 we arrive at the quantum corrected collisionless nonparabolic hydrodynamic transport model in one dimension,

$$\frac{\partial n}{\partial t} = -\frac{\partial}{\partial x}(nv) \qquad (12)$$

$$\frac{\partial v}{\partial t} = -v\frac{\partial v}{\partial x} - \frac{1}{m^*\mu}\frac{\partial}{\partial x}(U_{eff}+\frac{Q}{3}) - \frac{2}{3nm^*}\frac{\partial}{\partial x}[nv(w-w_q-\frac{m^*}{2}v^2] \qquad (13)$$

$$\frac{\partial w}{\partial t} = -v\frac{\partial w}{\partial x} - v\frac{\partial}{\partial x}(U_{eff}+\frac{Q}{3}) - \frac{2}{3n}\frac{\partial}{\partial x}[nv(\omega+\varphi)(w-w_q-\frac{m^*}{2}v^2]-2w_q\frac{\partial v}{\partial x} \qquad (14)$$

with nonparabolic correction terms $\mu(n,v,w) = 1 + 2\alpha\omega(v,w)(w-w_q-\frac{m^*}{2}v^2)$, $\nu(n,v,w) = \frac{\omega(v,w)}{\mu(n,v,w)}$, $\omega(v,w) = \frac{1}{1+\alpha m^*v^2}$, $\varphi(n,v,w) = \frac{10\alpha}{3}(\frac{\omega(v,w)}{\mu(n,v,w)})^2(w-w_q-\frac{m^*}{2}v^2)$ and quantum corrections $w_q = -\frac{\hbar^2}{24m^*}\frac{\partial}{\partial x}(\frac{1}{n}\frac{\partial n}{\partial x})$ and $Q = -\frac{\hbar^2}{2m^*}\frac{1}{\sqrt{n}}\frac{\partial^2\sqrt{n}}{\partial x^2}$. Q is the standard form of the quantum potential and w_q is a purely quantum mechanical contribution as first noted by Wigner.

4. Results and Conclusions

An initial study has been made of a double-barrier heterostructure, $Al_{x_c}Ga_{1-x_c}As$ barriers with $x_c = 0.3$, doped to $10^{14}cm^{-3}$ with source and drain regions of length

200 Å and doped to 10^{16}cm^{-3}. This structure is shown in Figure 2 with the source and drain regions excluded. In the self-consistent simulations, the heterostructure was treated using the Anderson rule; $U_{eff} = -q\Psi_{eff} - \chi(x)$, where $\chi(x)$ $(\Delta E_c = 0.697x_c)$ is a position dependent electron affinity and the applied potential Ψ_{eff} was determined from solving Poisson's equation. The dissipation mechanisms were treated in a classical manner [7] and we assumed $T_{lattice} = 300K$. The profile result of w_e and w_q, parameters intrinsic to the energy balance equation, are displayed in Figure 1 for $\Psi_{applied} = 8mV$. Figure 2 shows the corresponding profile result for Q along with the conduction band E_c $(E_c = U_{eff})$. This study is the first to use the three quantum hydrodynamic (QHD) equations to study transport in heterostructure devices. The key additional feature of the QHD equations is the incorporation of the effects of density gradients through the quantum potential. Insofar as a classical solution does not exist in the presence of barriers, Q must at least cancel the effects of the barriers and permit carrier transport. Indeed $\frac{Q}{3}$ does approximately balance the barrier potential a result consistent with [2].
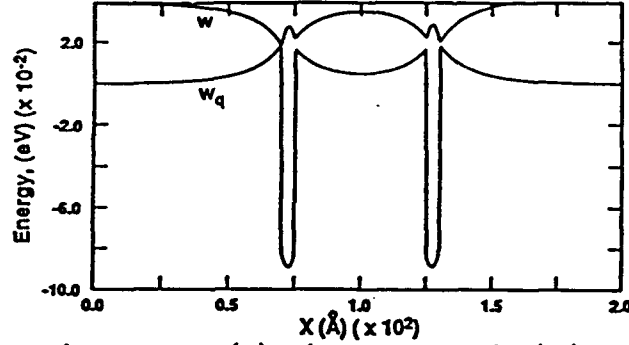
Figure 1. Average energy (w) and quantum correction (w_q) at $\Psi_{applied} = 8mV$.
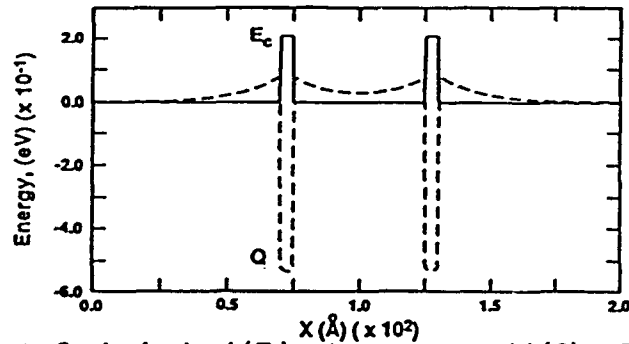
Figure 2. Conduction band (E_c) and quantum potential (Q) at $\Psi_{applied} = 8mV$.

1. Hot Carriers in Semiconductors, in *Solid-State Electronics* 32, 12, 1989.
2. H.L. Grubin and J.P. Kreskovsky *Solid-State Electronics* 32, 12, p. 1071, 1989.
3. D.L. Woolard, M.A. Littlejohn and R.J. Trew to be published
4. E.M. Azoff *IEEE Trans. Electron Devices*, 36, 4, p. 609, 1989.
5. M.A. Stroscio *Superlattices and Microstructures*, 2, 1, p. 83, 1986.
6. M.G. Ancona and G.J. Iafrat. *Physical Review* B, 39, 10, p. 9536, 1989.
7. D.L. Woolard et. al. *Solid-State Electronics* 32, 12, p. 1347, 1989.

COMPARATIVE NUMERICAL SIMULATIONS OF A GaAs SUBMICRON FET USING THE MOMENTS OF THE BOLTZMANN TRANSPORT AND MONTE CARLO METHODS

J. P. Kreskovsky, G. A. Andrews, B. J. Morrison and H. L. Grubin
Scientific Research Associates, Inc.
P. O. Box 1058
Glastonbury, Connecticut 06033
(203) 659-0333

Abstract

Numerical simulations of a submicron GaAs FET have been performed using both Monte Carlo (MC) methods and the Moments of the Boltzmann Transport Equation (MBTE). The I-V characteristics as well as details of the internal distribution of carriers and potential were obtained. The MC calculations show no regions or negative forward conductance. However, the MBTE results show that negative forward conductance can be present or absent depending on the value of the thermal conductivity.

Introduction

Recently, both MC and MBTE or hydrodynamic approaches to device simulation have been receiving greater attention. This is a result of the need to investigate, and accurately simulate, electron dynamics in submicron devices. With this reduction in device size nonequilibrium effects neglected in the more conventional drift and diffusion approach must be considered. This gives impetus to the use of the more complex but more fundamental methods mentioned above.

While attention has focused on the implementation of MC and MBTE simulation procedures, a direct comparison of these methods, applied to the same problem, has not been made. This paper presents the results of such a comparative study. While the results are preliminary in nature, they do provide significant insight and direction for future work.

Analysis

The device considered is a simple GaAs FET with a 0.6 micron source-drain spacing and a symmetrically placed 0.2 micron gate. The device depth is taken as 0.1 micron with a uniform doping of $1 \times 10^{17}/cm^3$.

The MC procedure employed follows closely that outlined by Hockney and Eastwood [1]. The change in particle momentum and position are advanced in time under the influence of an applied field assuming a non-parabolic band structure. At the end of the free flight period, the particles are scattered. This process is repeated for all particles until a specific increment in time is reached. At that time a space charge distribution is computed from the particle distribution using a cloud-in-cell algorithm and Poisson's equation is solved updating the potential. This process is repeated until a statistical steady state is achieved. The MC procedure considers transfer the between Γ, X and L valleys.

The MBTE approach solves continuity, momentum and energy equations for Γ and L valley electrons. These equations are expressed as

(1) $\quad \partial n_1/\partial t = -\partial n_1 V/\partial x - n_1/\tau_1 + n_2/\tau_2$

$$(2) \quad \partial n_1 V_1 / \partial t = - \partial n_1 V_1 V_1 / \partial x - 1/m \ \partial n_1 T_1 k_o / \partial x +$$

$$n_1 e / m_1 \ \partial \phi / \partial x + \mu \ \partial^2 V / \partial x^2 - n_1 V / \tau_3$$

$$(3) \quad \partial n_1 T_1 / \partial t = - \partial n_1 V_1 T_1 / \partial x - 2/3 \ n_1 T_1 \ \partial V / \partial x +$$

$$2/3 K_o \ \partial / \partial x \ (\kappa \partial T_1 / \partial x) - n_1 T_1 / \tau_5 + n_2 T_2 / \tau_6 +$$

$$m_1 V_1^2 / 3 k_0 \ (n_2 / \tau_2 - n_1 / \tau_1) + 2 n_1 m_1 V_1^2 / 3 k_0 \ \tau_3$$

for Γ (subscript = 1) valley carriers. A similar system with subscript 2 applies to the L valley carriers. The equations are expressed here in one-dimensional form. The extension the two-dimensional case considered here is straight forward. In Eq. (2), the momentum equation, the coefficient of the fourth term on the RHS, μ, controls dissipation. This term arise from the nonspherical nature of the distribution function. In general, we assume this coefficient to be zero or very small. If non-zero and significant, the slope of the I-V characteristics at low fields will be significantly in error and below that anticipated from the low field mobility of the material in question. The coefficient, κ, appearing in the third term on the RHS of the energy equation, is the thermal conductivity of the electron gas. While its precise value is unknown, it is important and its role in the behavior of the solutions is investigated here. If it is zero, there will be no heat transfer by conduction while if it approaches infinity the energy equation reduces to

$$(4) \quad \nabla^2 T = 0$$

The quantities indicated by τ in Eqs. (1-3) are relaxation times determined from a prior evaluation of the collision integrals ame are taken as energy dependent in the present study [2].

Equations (1-3) and the L valley counter parts are coupled to Poissons' equation and the resulting system is solved using an algorithm based on the LBI (Linearized, Block, Implicit) technique developed by Briley and McDonald [3] and paralleling that used by Kreskovsky and Grubin [4] for solving the drift and diffusion equations. The algorithm is highly efficient and a high degree of parallelism is present making it ideally suited for vector processing. A brief discussion of its application to the present system may be found in [5].

Results

Simulations of the FET shown in Fig. 1 were performed for gate bias levels of 0.1, 0.3 and 0.5 volts with the drain bias varying from 0 to 2.0 volts. Both the MC and the MBTE simulations used an 11 x 61 grid point mesh yielding square cells with 100Å spacing. Three ·'ousand (3000) particles (an average of 5 per cell) were used in the MC calculations. The MBTE simulations required approximately 2.5 minutes of Cray XMP CPU time per bias point. The MC simulations required approximately 6 minutes per bias point. Results for MC simulations using 9000 particles showed no significant differences but required significantly greater CPU time.

Figure 2 shows the I-V characteristics of the device as determined from the two simulation approaches. At a gate bias of 0.1 volts it is observed that both the MC and MBTE results show good agreement. At this bias level, the Γ valley electrons do not become too hot (maximum local temperature of 1320° K) and electron transfer is a maximum of approximately 52% Γ valley and 38% L valley. As the gate bias is increased to 0.3 volts we see that the agreement is good up to approximately 0.6 volts V_{ds}. Above this value of drain bias the MBTE result :xhibits negative forward conductance. No such effect is observed for the MC result. At V_{ds} = 2.0 volts, the maximum Γ valley electron temperature from the MBTE result reaches 2400° K and the population of carriers in the Γ and L valleys is 20% and 80% respectively at the drain; the point of maximum transfer.

The trend of wider disagreement between MC and MBTE predictions is seen to grow as the gate bias is increased. At V_{gs} = 0.5 volts we observe that negative conductance appears in the MBTE simulation at about V_{ds} = 0.5 volts. Here the Γ valley temperatures reaches a maximum of 2925° K and at the drain the population of carriers in the Γ valley is only 10% of the total. At V_{ds} = 1.0 volts, the maximum Γ valley temperature is only 1860° K, and the Γ valley population at the drain is 25%. It should be noted the population of carrier in the X valley from the MC results was negligible at all bias levels.

The reason for the negative conductance is straight forward. Once the bias levels become high enough that sufficient energy is supplied to the carriers to cause them to transfer, the possibility exists that the carriers will transfer at a rate which is greater than the Γ valley velocity increases. Since the current is predominantly through Γ valley transport, if this occurs the current will decrease. This is precisely what is happening here and further suggests two possible sources of the discrepancy. First, the scattering rate constants, τ_{ij}, which are functions of the electron temperature, could have an inappropriate high temperature variation. However, under uniform field conditions the results obtained from both MC and MBTE simulations yield reasonable velocity field curves which are in general agreement with experimental data. This suggests that the problem lies elsewhere, and is associated with spatial nonuniformities in the solution. Examination of the continuity and momentum equations does not reveal any possible sources of the problem. (As discussed earlier, the coefficient of the dissipation term, μ, affects the low field portion of the I-V curve, but has little effect at high bias). In the energy equation, the only term which can be regarded with suspicion is the heat conduction term. This prompted us to examine the result if the thermal conductivity were infinite. This result is shown for V_{gs} = 0.5 volts as the curve labeled "B" in Fig. 2 and shows that indeed such a limit removes the negative conductance. In fact, we have been able to obtain results which match the MC results by varying the value of κ. However, the required value of κ is highly bias dependent. For example, at V_{gs} = 0.5 and V_{ds} = 1.0 a value of κ five times that used in the generations of the curve labeled "A" gave the MC result, yet at V_{ds} = 2.0 a value of κ 100 times the curve "A" value gave a current that was still 20% below the MC result. One other possibility is that the scattering rates should be based on a historically averaged temperature, rather than the local temperature as done at present. Clearly, more research needs to be done here.

Finally, we compare the distribution of potential and carrier density within the device, as determined from the two procedures, in Figs. 3 and 4. Fig. 3 shows the potential distribution at V_{gs} = 0.3 and V_{ds} = 1.0. Here the I-V curves are in good agreement, and in Fig. 3 we observe that the potential distributions are almost identical. The MC result shows a slight amount of noise which is to be expected since the space charge must be derived from the discrete particle distribution. However, the similarity in the results indicates similarity in the space charge distribution. Fig. 4 then shows a comparison of the MC particle distribution (lower half) with the density contours obtained from the MBTE simulation. Here the gate lies along the centerline of the figure. The similarity in the depletion region is evident here.

Conclusions

A comparative study of a submicron GaAs FET has been performed using MC and MBTE approaches. The two approaches were found to give good agreement in the I-V characteristics at low values of gate bias. As the gate bias was increased, the MBTE result exhibited negative forward conductance whereas only positive forward conductance was observed in the MC results. The presents of negative conductance was found to be dependent on the magnitude of the thermal conductivity coefficient in the Γ valley energy equation, with negative conductance disappearing altogether as the thermal conductivity approached infinity. At bias levels where good agreement was present in the predicted I-V

144

result, good agreement was also observed in the potential and carrier distributions within the device. Overall the results indicate that both MC and MBTE approaches are viable for submicron device simulation, however, further research is needed to provide proper closure of the MBTE approach, specifically with regard to the role of the heat flux terms in the energy equation, and the evaluation of scattering rates based on historic rather than local quantities.

Acknowledgements

References

1. R. W. Hockney and J. W. Eastwood, Computer Simulation Using Particles, Adam Hilger, (1988).
2. H. L. Grubin, D. K. Ferry, G. J. Iafrate and J. R. Barker, VLSI Electronics, 3, (1982).
3. W. R. Briley and H. McDonald, J. Comp. Phys. (1980).
4. J. P. Kreskovsky and H. L. Grubin, J. Comp. Phys. (1987).
5. J. P. Kreskovsky, M. Meyyappan and H. L. Grubin, NUMOS I, (1986).
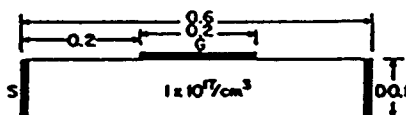
Figure 1. Schematic of GaAs FET

A= FINITE THERMAL CONDUCT
B= INFINITE THERMAL COND.

Figure 2. Comparison of I-V Characteristics as determined from Moment Equations and Monte Carlo Solutions for a GaAs FET.

(a)   (b)

Figure 3. Comparison of Potential Surface a) Moment Equations b) Monte Carlo Simulation. $V_{gs} = 0.3$, $V_{ds} = 1.0$.

SOURCE   DRAIN
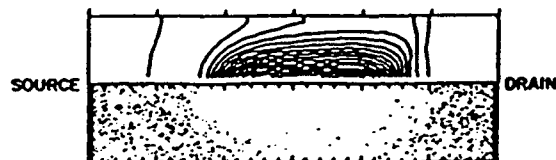
Figure 4. Comparison of Monte Carlo Particle Distribution (lower) with Moment Equation Solution Density Contours (upper) for a GaAs FET. $V_{gs} = 0.3$, $V_{ds} = 1.0$.

Fundamental Research on the Numerical Modelling of Semiconductor Devices and Processes

Papers from NUMOS I, the First International
Workshop on the Numerical Modelling of Semiconductors
11th - 12th December 1986, Los Angeles, USA.

Edited By
J.J.H. Miller
Numerical Analysis Group, Trinity College, Dublin

BOOLE

BOOLE PRESS

THE MOMENTS OF THE BOLTZMANN TRANSPORT EQUATION AS APPLIED TO THE GALLIUM ARSENIDE PERMEABLE BASE TRANSISTOR

J.P. KRESKOVSKY, M. MEYYAPPAN and H.L. GRUBIN
Scientific Research Associates, Inc., Glastonbury, CT 06033, USA

ABSTRACT: Solutions to the first three moments of the Boltzmann transport equation and Poisson's equation are obtained for a permeable base transistor (PBT) using linearized, block implicit (LBI) and ADI techniques. Two level electron transfer is considered. The results of the simulations are compared to results obtained from the drift and diffusion equations. The comparison indicates that nonequilibrium transport and velocity overshoot are important in the PBT. The predicted *I-V* characteristics of the device show substantially higher current levels and a higher cutoff frequency are obtained with the moment equations.

INTRODUCTION

Numerical and experimental studies of the PBT were first performed by Bozler and Alley [1]. Their results suggested that a PBT could obtain cutoff frequencies as high as 300 GHz. This result has not yet been obtained. In some of their drift and diffusion computations Bozler and Alley [1] used a two piece velocity-field curve in an attempt to account for nonequilibrium effects. The velocity varied linearly with field up to a value of 4 kV/cm after which it was held constant at 2×10^7 cm/sec. This approach will yield higher current levels, but it offers no improvements in either the qualitative or quantitative consequences of nonequilibrium transport. Nonequilibrium effects, in the presence of nonuniform fields, result in carrier and velocity distributions that are significantly different than those offered by excess saturation velocity coupled to the drift and diffusion equations.

Nonequilibrium effects in a PBT have been investigated using Monte Carlo methods by Hwang et al. [2]. Their results showed significant differences compared to the drift and diffusion results; the peak velocity along the center of the channel exceeded 4×10^7 cm/sec and the density distribution was also significantly different. The Monte Carlo results also indicated a cutoff frequency approximately 60% higher than the corresponding drift and diffusion prediction.

In the present note, the performance of the PBT structure shown in Fig. 1 is analyzed using both the drift and diffusion and the moments of the Boltzmann transport equation, (MBTE). The base penetration of the present device structure is only one half of the channel width. This in itself results in a higher f_T, compared to devices with equal base penetration and channel width [3]. However, further and substantial increases in f_T are observed from the MBTE predictions as well as higher current levels.
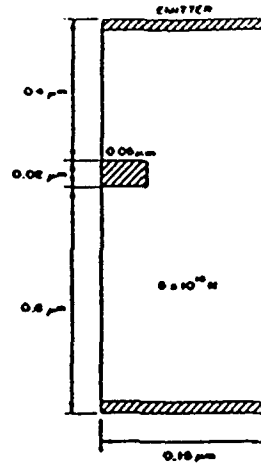
Figure 1: Dimensions and doping level of the simulated PBT.

ANALYSIS AND NUMERICAL METHOD

The equations to be solved are obtained by taking the first three moments of the Boltzmann transport equation. The resulting equations represent conservation of mass, momentum, and energy, and may be expressed for the central valley carriers as

$$\frac{\partial n_1}{\partial t} = - \frac{\partial n_1 V_1^i}{\partial x_i} - n_1 \Gamma_{11} + n_2 \Gamma_{12} \tag{1}$$

$$\frac{\partial n_1 V_1^i}{\partial t} = - \frac{\partial n_1 V_1^i V_1^i}{\partial x_j} - \frac{k_0}{m_1} \frac{\partial n_1 T_1}{\partial x_i} - \frac{1}{m_1} \frac{\partial \sigma_1^{ij}}{\partial x_j} + \frac{e n_1}{m_1} \frac{\partial \psi}{\partial x_i} - n_1 V_1^i \Gamma_{13} \tag{2}$$

$$\frac{\partial n_1 T_1}{\partial t} = - \frac{\partial n_1 V_1^j T_1}{\partial x_j} - \frac{5}{3} n_1 T_1 \frac{\partial V_1^j}{\partial x_j} - \frac{2}{3} \frac{\sigma^{ij}}{k_0} \frac{\partial V^i}{\partial x_j} + \frac{2}{3 k_0} \frac{\partial^x \partial T_1 / \partial x_j}{\partial x_j}$$

$$- n_1 T_1 \Gamma_{14} + n_2 T_2 \Gamma_{15} + \frac{m_1}{3 k_0} (2 n_1 \Gamma_{13} - n_1 \Gamma_{11} + n_2 \Gamma_{12}). \tag{3}$$

A similar system is obtained for the satellite valley carriers. These equations are coupled to Poisson's equation

$$\epsilon \nabla^2 \psi = e(n_1 + n_2 - N_D). \tag{4}$$

In eqs. (1)–(4) n is carrier density, V^i the velocity vector, T the carrier temperature, ψ the electrostatic potential, k_0 is Boltzmann's constant, m is the effective electron mass. Γ_{ij} are scattering rate constants, σ^{ij} is a stress tensor arising from the nonspherical nature of the distribution function, and κ is a thermal conductivity. The scattering rates are determined a priori from evaluation of the collision integrals [4].

Equations (1)–(3), their satellite valley counterparts, and Poisson's equation form a coupled system of nine nonlinear PDE's in two space dimensions of the form

$$\frac{\partial H(\phi)}{\partial t} = D(\phi) + S(\phi) \qquad (5)$$

where ϕ represents the vector of unknowns, H and S represent nonlinear functions of ϕ, and $D(\phi)$ represents a nonlinear, partial differential operator. To solve this system Poisson's equation is decoupled from the remaining eight equations by differencing the electric field appearing in the momentum equation at the explicit time level. The remaining equations are solved by direct application of linearized block implicit (LBI) techniques [5]. The equations are first time differenced

$$\frac{\Delta H(\phi)}{\Delta t} = \beta[D(\phi)^{n+1} + S(\phi)^{n+1}] + (1 - \beta)[D(\phi)^{n} + S(\phi)^{n}] . \qquad (6)$$

The nonlinear operators are then linearized using Taylor series, as for example:

$$D(\phi^{n+1}) = D(\phi^{n}) + \frac{\partial D}{\partial \phi}\bigg|^{n} \Delta \phi^{n+1} + O(\Delta t^{2}) . \qquad (7)$$

Equation (6) may then be expressed as

$$(A + \Delta t L)\Delta \phi^{n+1} = \Delta t[D(\phi^{n}) + S(\phi^{n})] \qquad (8)$$

where

$$A \equiv \left(I - \beta \Delta t \, \frac{\partial S}{\partial \phi}\bigg|^{n}\right) \qquad (9a)$$

$$L \equiv -\beta \, \frac{\partial D}{\partial \phi}\bigg|^{n} . \qquad (9b)$$

Equation (8) is then split following the ADI procedure of Douglas and Gunn [6].

$$(A + \Delta t L_{x})\Delta \phi^{*} = \Delta t[D(\phi^{n}) + S(\phi^{n})] \qquad (10a)$$

$$(A + \Delta t L_{y})\Delta \phi^{**} = A\Delta \phi^{*} \qquad (10b)$$

and

$$\Delta \phi^{n+1} = \Delta \phi^{**} + O(\Delta t^{2}) . \qquad (11)$$

As a result of the ADI splitting, the number of operations required to solve the remaining coupled equations varies linearly with the total number of grid points, and due to the linearization process, the solution can be advanced in time without introducing nonlinear iteration. After solution of the continuity, momentum, and energy equations, Poisson's equation is solved using a scalar ADI procedure to complete a time step. Steady solutions, such as those to be presented here, are obtained from the long time asymptotic transient solution. Since transient accuracy is not of interest under such conditions, the time step can be spatially scaled to speed convergence to steady state.

The solutions to the drift and diffusion equations reported here were obtained using a related procedure, described in detail in [7].

COMPUTED RESULTS

The computed I-V characteristics of the device shown in Fig. 1 are presented in Fig. 2. The results for the MBTE calculations were extrapolated to the origin, as indicated by the long broken lines. The shorter dashed curves show the results for the DDE [3]. The comparison shows that the predicted current levels are significantly higher for the MBTE solutions, a result consistent with FET calculations performed by the present authors, as well as by Cook and Frey [8] who used a highly simplified momentum-energy transport model. The present MBTE results also indicate a region of negative differential forward conductivity at $V_{BE} = 0.6$ V. The origin of this phenomena is believed to be a consequence of electron transfer. The presence of a dc negative forward conductance is also a feature of PBT measurements [9], but is clearly absent from DDE simulations.

A comparison between the total carrier density distribution along the center of the channel for drift and diffusion and MBTE solutions is shown in Fig. 3 for $V_{CE} = 1.0$ V and $V_{BE} = 0.4$ V. The MBTE prediction for the gamma valley carrier density is also shown. The drift and diffusion result was obtained using an equilibrium velocity-field curve, and yields results qualitatively similar to that of Bozler and Alley.

As seen in Fig. 3 for the DDE simulations, the carrier density reaches a maximum between base contacts. Here, with the velocity in saturation and the cross sectional area at a minimum, the carrier density must increase to maintain current continuity. In the MBTE simulation the constraints of current continuity are more complex. First a decrease in the cross sectional areas is, as in the DDE, accompanied by an increase in field along the channel. The field increase under both equilibrium and nonequilibrium conditions is qualitatively similar, as may be observed from Fig. 4 which shows the potential distribution along the center of the PBT channel.
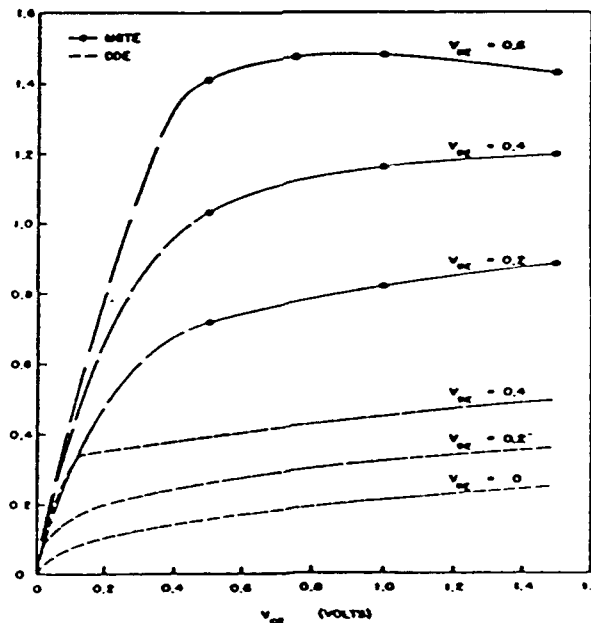
Figure 2: Collector current vs. collector emitter voltage for different values of base emitter voltage.
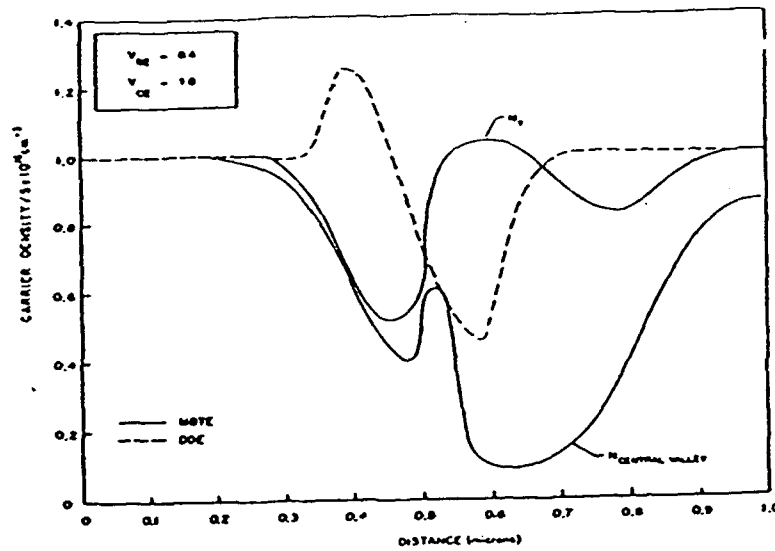
Figure 3: Carrier density vs. distance along center of channel for the PBT.

However, consequent changes in electron temperature, both increasing and decreasing, lag behind the equilibrium state. This leads to velocity overshoot and a delay in electron transfer. As a result, for nearly the first half of the device transport is almost exclusively gamma valley transport. The implication is that if the gamma valley carrier velocity increases with increasing field, then the product of density and cross sectional area normal to current flow must decrease to maintain current continuity. Since the velocity increases faster than the area decrease the carrier density decreases.

At moderate bias levels typical FET calculations show a decreasing field as the gate region is passed. This also occurs in the PBT. Now, as the cross sectional area increases the gamma valley carriers exhibit a decrease in velocity. It must be noted, however, that for the parameters of the calculations the L valley carriers make a negligible contribution to current. Thus a decrease in carrier velocity results in a net
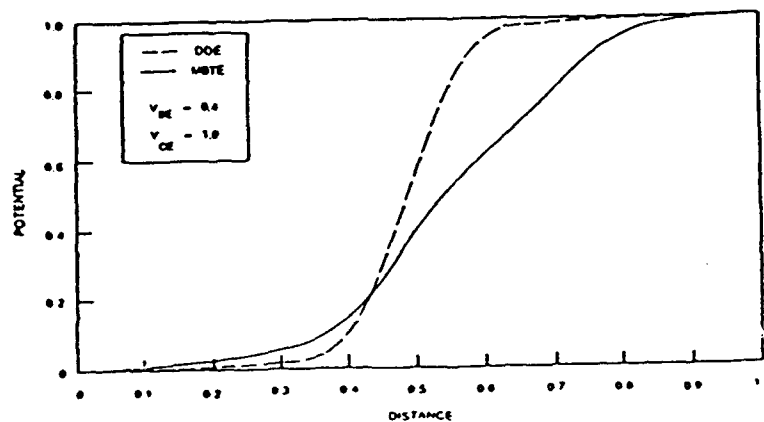
Figure 4: Potential vs. distance along center of channel for the PBT.
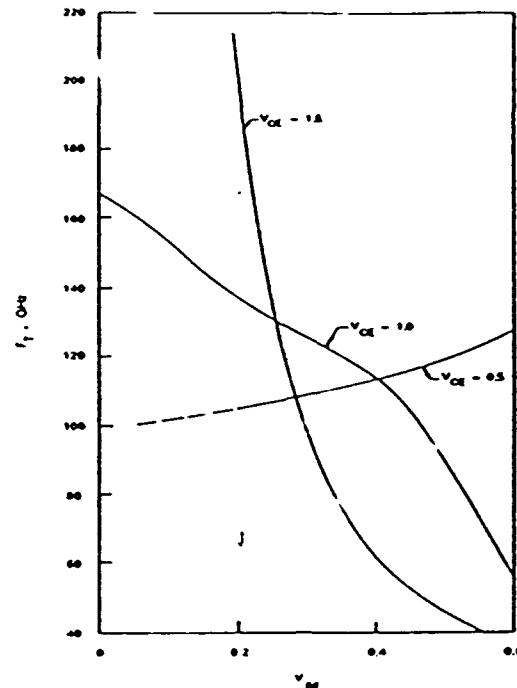
Figure 5: Cutoff frequency vs. base emitter voltage.

increase in carrier concentration. However initially, the decrease in field is not accompanied by a corresponding temperature decrease (as experienced in the uniform field calculations). Thus, the high gamma valley temperature results in transfer to the L valley giving rise to the second minimum in the gamma valley carrier density shown in Fig. 3. Further toward the drain, the field decreases. However, relaxation is incomplete and the field at the collector is not equal to the field at the emitter. Also note, the MBTE potential distribution gives rise to a slightly higher field upstream of the base, and a lower field, over a longer distance, downstream of the base compared to the DDE result. More significantly the electron temperature at the collector exceeds that at the emitter. It is noted that as the field relaxes, the electrons transfer back to the central valley.

All of the results presented thus far are qualitatively similar to the Monte Carlo results of Hwang et al. [2]. However, the results for cutoff frequency, shown in Fig. 5, are qualitatively different. At $V_{CE} = 0.5$ V, f_T is seen to increase with V_{CE} from approximately 100 GHz to 130 GHz, value 2.5 to 3 times greater than that obtained using the DDE. However, for higher values of V_{CE}, f_T is shown to decrease with increasing V_{BE}. For $V_{CE} = 1.5$ V, an f_T in excess of 200 GHz is predicted for $V_{BE} + 0.2$ V; while at $V_{BE} = 0.6$ V, f_T has dropped below 40 GHz, a result that is qualitatively consistent with the presence of the negative forward conductance, and one that would find no explanation in the DDE calc tions.

ACKNOWLEDGEMENTS

REFERENCES

[1] C.O. Bozler and G.D. Alley, IEEE Trans. Electron Dev. ED-27(6) (1980).

[2] C.-G. Hwang, D.H. Navon and T.-W. Tang, IEEE Electron Dev. Lett. EDL-6(3) (1985).

[3] M. Meyyappan, J.P. Kreskovsky and H.L. Grub'n, SRA report, in preparation.

[4] H.L. Grubin, D.K. Ferry, G.J. Iafrate and J.R. Barker, VLSI Electronics 3 (1982).

[5] W.R. Briley and H. McDonald, J. Comp. Phys. (1980).

[6] J. Douglas and J.E. Gunn, Num. Math. 6 (1964).

[7] J.P. Kreskovsky and H.L. Grubin, J. Comp. Phys. (1987).

[8] R.F. Çook and J. Frey, COMPEL 1(2) (1982).

[9] See, e.g., C.O. Bozler, M.A. Hollis, K.B. Nichols, S. Rabe, A. Vera and C.L. Chen, IEEE Electron. Dev. Lett. EDL-6(9) (1985).

Electron velocity overshoot and valley repopulation effects in diamond

M. A. Osman, H. L. Grubin, and J. P. Kreskovsky

Scientific Research Associates, Inc., P. O. Box 1058, Glastonbury, Connecticut 06033

Using ensemble Monte Carlo procedures, electron velocity overshoot and transient valley repopulation in diamond have been investigated as a function of field strength and orientation. It is found that the response of electrons to the sudden application of a 50 kV/cm electric field along $\langle 100 \rangle$ results in a velocity transient with a maximum of 1.9×10^7 cm/s after 200 fs and a steady state value of 1.30×10^7 cm/s. For a field along $\langle 110 \rangle$, the corresponding maximum and steady state values are 2.0×10^7 and 1.4×10^7 cm/s, respectively. The calculated temporal and spatial duration of velocity overshoot in diamond is longer than that of silicon but shorter than that of gallium arsenide.

The superior thermal conducting characteristics of semiconducting diamond, coupled to hole and electron mobilities and saturated drift velocities that are superior to silicon, make it an ideal candidate for high-speed, high-power, and radiation-resistant device applications.[1-3] To exploit this, much effort is under way in the growth and doping of synthetic diamond, while interest in natural diamond remains high.[4-7] The successful development of a semiconducting diamond device technology requires the determination of a number of key transport parameters, such as mobilities, transient velocities, etc. The mobility and steady state drift velocity of electrons and holes in natural diamond have been investigated by Nava *et al.*[2] and Reggiani *et al.*[3] In this letter we examine two effects occurring in diamond for the first time: (1) velocity overshoot and (2) transient valley repopulation (as have been investigated in Si and Ge[8]). The calculations reported below are for electrons subjected to sudden changes of field with values 20 and 50 kV/cm. Additionally, for fields of 50 kV/cm, the orientation dependence is studied. It is found that velocity overshoot occurs on temporal scales of the order of 400 fs, which is longer than that of silicon and the same order as that of gallium arsenide. The spatial duration of the overshoot is of the order of 400 Å, less than that of gallium arsenide[9] but larger than that of silicon.[10] Thus nonequilibrium effects are expected to be important at feature sizes below 400 Å.

In the ensemble Monte Carlo procedure used in this study of diamond, the kinetics of electrons are determined relative to the minimum in k space at which the valley is located. Electrons located in valleys on opposite sides of the same axis, e.g., the (100) and $(\bar{1}00)$ valleys, are not distinguishable, and symmetry allows specific calculations to be performed for three ellipsoidal valleys ($m_l = 1.4$, $m_t = 0.36$) located along the $\langle 100 \rangle$, $\langle 010 \rangle$, and $\langle 001 \rangle$ crystallographic orientations. The Monte Carlo procedure takes into account scattering by elastic intravalley acoustic phonons, g-type intervalley phonon scattering between two parallel valleys, and f-type intervalley phonon scattering between four perpendicular valleys. Following an f-type scattering, the final valley is selected from the possible perpendicular valleys using a random number. The interaction between the electrons and ionized impurities is ignored in this investigation. The longitudinal-optical phonon modes assist intervalley g scattering and have a temperature of 1900 K, while

longitudinal-acoustic and transverse-optical phonon modes assist intervalley f scattering and have temperatures of 1500 and 1720 K, respectively.[8] These energies of the intervalley phonons are considerably larger than the corresponding values in Si.[8] A single value, 8×10^8 eV/cm, was used for all of the intervalley deformation potentials, and a value of 8.7 eV was used for the intravalley acoustic deformation potential.[2] Initially, the electrons were assumed to be in thermal equilibrium corresponding to a Maxwellian distribution at room
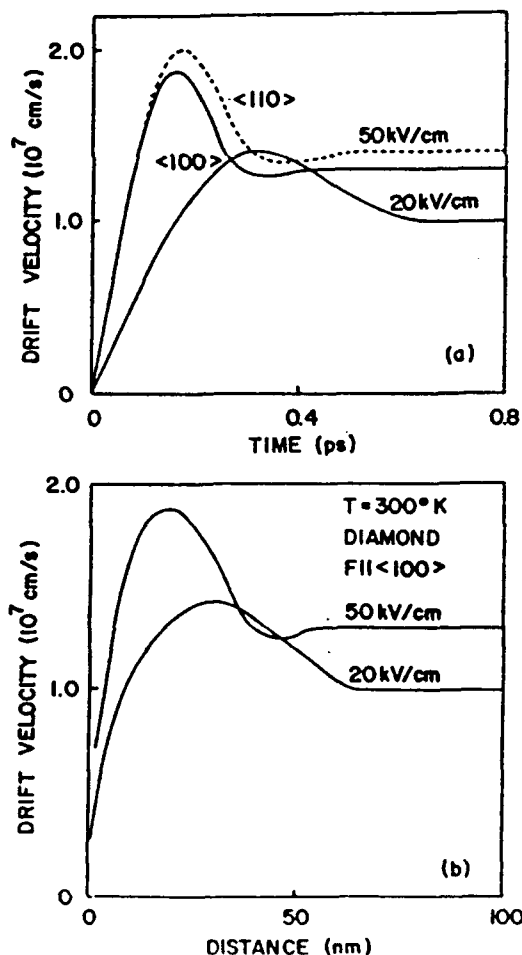

FIG. 1. Drift velocity of electrons in diamond as a function of (a) time and (b) position at 300 K. The orientations of the electric field are shown for each curve.

temperature. The electric field was then switched on and the time evolution of the electron ensemble was examined for fields of 20 and 50 kV/cm applied along the $\langle 100 \rangle$ direction. To illustrate the orientation dependence of valley repopulation and velocity overshoot, the calculations were repeated for a 50 kV/cm electric field along the $\langle 110 \rangle$ direction.

The time evolution of the electron velocity during the first picosecond after switching on the electric field is shown in Fig. 1. For a 50 kV/cm electric field along the $\langle 100 \rangle$ direction, the transient velocity reaches a maximum of 1.9×10^7 cm/s after 200 fs and then decreases to a steady state value of 1.30×10^7 cm/s. When a 20 kV/cm field is applied along $\langle 100 \rangle$, the velocity increases more gradually to a maximum of 1.4×10^7 cm/s after 300 fs, and then decreases to a steady value of 1.0×10^7 cm/s. Changing the orientation of the electric field at 50 kV/cm to $\langle 110 \rangle$, increases the maximum transient velocity and steady state velocity to 2.00×10^7 and 1.4×10^7 cm/s, respectively. The steady state results are consistent with the earlier experimental work reported by Nava et al., which showed higher drift velocities in natural diamond for fields along $\langle 110 \rangle$ compared to that along $\langle 100 \rangle$.[2] In all of the above situations, the maximum velocity is about 50% higher than the corresponding steady state velocity. The magnitudes of the maximum and the steady state velocities are larger than those reported for Si at 50 kV/cm.[8] Additionally, the time duration of velocity overshoot is longer in diamond than in Si. For example, for a 20 kV/cm field along $\langle 100 \rangle$, the time duration of overshoot is approximately 250 fs in Si[8] compared to 400 fs in diamond as can be seen from Fig. 1(a). *The longer overshoot duration makes it possible to design diamond devices with larger dimensions than Si devices, while maintaining shorter transit times across the device.* Figure 1(b) shows that the distance over which overshoot effects are pronounced is approximately 400 Å for diamond. This overshoot distance is longer than the 300 Å in silicon at comparable field values.[10]

Figure 2 displays the time evolution of the average energy of the carriers subject to fields of 20 and 50 kV/cm. For the 50 kV/cm field, and during the first 200 fs, the average energy of the carriers rises rapidly and is nearly independent of orientation. However, at longer times the average energy for the field along $\langle 110 \rangle$ is slightly higher than that along the $\langle 100 \rangle$ direction. In response to a 20 kV/cm field along $\langle 100 \rangle$, the average energy increases gradually during the first 400 fs and reaches a value of 130 mV which is below the threshold for the emission of g-type intervalley phonons. To illustrate the valley repopulation effects and how it influences the velocity overshoot, the mean energy and the number of electrons in individual valleys were determined. In Fig. 3, the average energy of the electrons in the (100) (dashed lines) and (001) valleys (solid lines) is plotted for 20 and 50 kV/cm electric fields. Note that, regardless of field orientation, e.g. along $\langle 100 \rangle$ or $\langle 110 \rangle$, the component of the electric field along the principal axis of the (001) valleys is zero, while it is finite along the minor axes where transport is determined by the smaller transverse effective mass. Consequently, the electric field heats the electrons in the (001) valleys at a rate faster than that of the (100) valleys. This
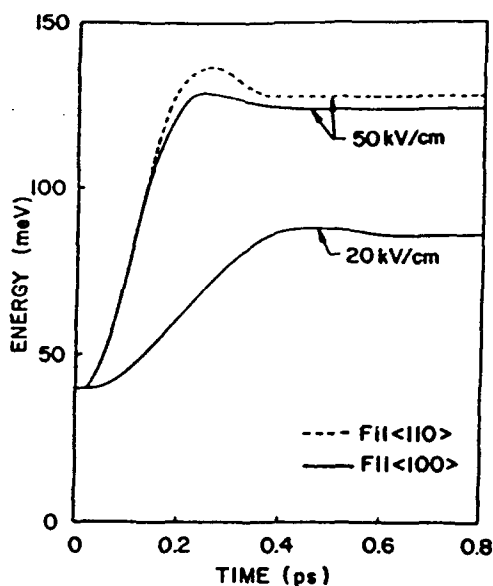

FIG. 2. Mean energy as a function of time for electrons in diamond. Different curves refer to different magnitudes and orientation of the electric field.

results in an initial rapid rise in average energy, as well as a higher average energy at longer times for the electrons in the (001) valleys relative to those in the (100) valleys (see Fig. 3 for both 20 and 50 kV/cm fields).

The (001) and (010) valleys are equivalent when the field is along $\langle 100 \rangle$ directions. Hence the rapid heating of electrons in these valleys results in a significant fraction of the electrons undergoing f-type intervalley phonon scattering to the (100) valleys, whose transport is governed by the larger longitudinal effective mass.[11] This reduces the population of the electrons in the (001) valleys while increasing that of (100) valleys, as can be seen in Fig. 4. For example, in response to a 50 kV/cm field along $\langle 100 \rangle$ [Fig. 4(a)], the population of the electrons in the (100) valleys rapidly in-

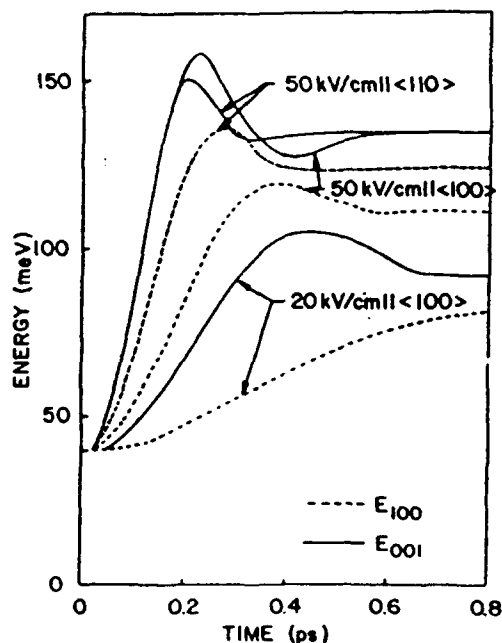FIG. 3. Mean energy as a function of time for electron in diamond in (001) valleys (solid lines) and (100) valleys (dashed lines) at 300 K.
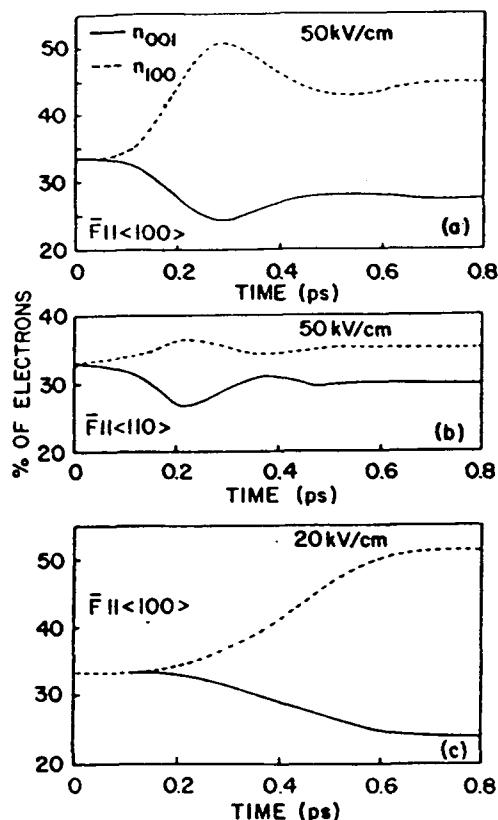
FIG. 4. Fraction of electrons in (100) valleys (dashed lines) and (001) valleys (solid lines) as a function time after switching on the field at 300 K: (a) $F = 50$ kV/cm $\|(100)$, (b) $F = 50$ kV/cm, $\|(110)$, and (c) $E = 60$ kV/cm $\|(100)$.

creases from its equilibrium value to 50% of the total after 300 fs and then gradually decreases to 45% after 700 fs. The physics behind this repopulation is displayed in Fig. 3, which shows that after 200 fs the average energy of the electrons in the (100) valleys exceeds the threshold for intervalley phonon emission. This leads to a larger fraction of f-type intervalley scattering, a reduction of the (100) valley population, and an increase in (001) valley population.

When the orientation of the field is changed to $\langle 110 \rangle$, the (100) and (010) valleys are equivalent; and transport in these valleys is through a combination of longitudinal and transverse effective masses. Several new features enter: First, the energy difference between the (001) and (100) valleys is smaller than for the field along $\langle 100 \rangle$; second, because of the contribution of the transverse effective mass, the electrons in the (100) valleys have a higher average energy and heat up at a faster rate compared to the situation when the field is along the $\langle 100 \rangle$ axis as can be seen in Fig. 3. Consequently, the difference between the f-type intervalley scattering from (100) to (001), and the intervalley scattering rate from (001) to (100) are smaller for the $\langle 110 \rangle$ orientation than for

the $\langle 100 \rangle$ orientation, and the population of the electrons in both (001) and (100) valleys changes by less than 5% from the initial thermal equilibrium values. The steady state energy of the electrons in the (100) valleys remains well below the f-type phonon emission threshold (see Fig. 3) when a 20 kV/cm is applied along the $\langle 100 \rangle$ axis, so that only a smaller fraction of electrons undergoes intervalley scattering to (010) and (001) valleys. Here the population of the electrons in the (100) valley gradually increases to 50% after 1 ps, while that of the (001) valleys decreases to 25% of the total electron population [Fig. 4(c)].

In conclusion, we have shown that the electrons in diamond exhibit orientation-dependent velocity overshoot and that the temporal and spatial duration of the overshoot is longer compared to Si. The maximum velocities were about 50% larger than the steady state velocities. The magnitudes of the maximum velocities following a 50 kV/cm field were 2.00×10^7 cm/s for $\langle 110 \rangle$ field orientation and 1.9×10^7 cm/s for $\langle 100 \rangle$ orientation. The valley repopulation is strongly affected by the orientation and the magnitude of the electric fields. Finally, the influence of the ionized impurity scattering on the above conclusions remains to be investigated. Because the known donor impurities, such as activated lithium, require more than 100 meV to ionize and release free electrons, the concentration of free electrons and ionized donors is strongly dependent on the electric field. This can influence the transient response of electrons to applied electric field due to the finite time required to release the electrons from the donor impurities.

[1]V. K. Baxhenov, I. M. Viokulin, and A. G. Gonor, Sov. Phys. Semicond. 19, 829 (1985).
[2]F. Nava, C. Canali, C. Jacoboni, L. Reggiani, and S. F. Kozlov, Solid State Commun. 33, 475 (1980).
[3]L. Reggiani, S. Bosi, C. Canali, F. Nava, and S. F. Kozlov, Phys. Rev. B 23, 3050 (1981).
[4]B. Singh, O. R. Mesker, A. W. Levine, and Y. Arie, Appl. Phys. Lett. 52, 1658 (1988).
[5]V. S. Vavilov, M. A. Gukasyan, M. I. Guseva, T. A. Karatygina, and E. A. Konorova, Sov. Phys. Semicond. 8, 471 (1974).
[6]K. L. Moazad, R. Nguyen, J. R. Zeidler, IEEE Electron Device Lett. EDL-9, 350 (1988).
[7]M. W. Geis, D. D. Rathman, D. J. Ehrlich, R. A. Murphy, and W. T. Lindley, IEEE Electron Device Lett. EDL-8, 341 (1987).
[8]C. Jacoboni and L. Reggiani, Rev. Mod. Phys. 55, 645 (1983).
[9]T. J. Maloney and J. Frey, J. Appl. Phys. 48, 781 (1977).
[10]D. K. Ferry and H. L. Grubin, Microelectron. J. 12, 5 (1981).
[11]K. Seeger, Semiconductor Physics (Springer, Vienna, 1973), Chap. 7.

Monte Carlo investigation of carrier-carrier interaction and ultrafast cooling of hot photoexcited carriers in GaAs

M.A. Osman, H.L. Grubin, and J.P. Kreskovsky
Scientific Research Associates, Inc.
P.O. Box 1058, Glastonbury, Connecticut 06033-6058

D.K. Ferry
Center for Solid State Electronics Research
Arizona State University, Tempe, Arizona 85281

Abstract

The role of the electron-electron (e-e), hole-hole (h-h), and electron-hole (e-h) interaction on ultrafast cooling of carriers in GaAs is examined for excess excitation energies of 40, 200, and 300 meV using an Ensemble Monte Carlo (EMC) approach. It is found that when the initial energy of the carrier is below the phonon emission threshold carrier-carrier (c-c) interactions stimulate either the optical phonon emission or absorption process depending on whether the initial energy of the carrier is above or below the thermal energy, respectively. The e-h interaction role is strong when excitation energy is below the LO phonon emission threshold.

Introduction

In recent years, the time-resolved femtosecond spectroscopy and optical time of flight measurement techniques have become an attractive tool for investigating the ultrafast dynamics of photoexcited carriers in semiconductors[1-3]. Considerable experimental information about the carrier-carrier and carrier-phonon interactions processes has been obtained. These processes show a strong dependence on carrier concentration, energy, electric field, temperature, and material properties. For example, at high electric fields the minority electrons in p-type InGaAs exhibit higher saturation velocities than majority carrier electrons[4], while in p-type Si they exhibit mobilities that are much lower than majority carrier electrons[5]. Similarly an increase in the electron energy-loss-rates in the presence of a cold hole plasma[6], and an absolute negative mobility for the minority electrons and holes at low temperatures in GaAs/AlGaAs quantum well structures have been observed[2]. In all these situations, the e-h interaction has been suggested as the possible responsible mechanism. More recently it has also been suggested that electrons excited by excess energy of 0.9 eV relax initially by emitting LO phonons and subsequently thermalize through e-e interactions[7]. These remarkable experimental results clearly demonstrate our limited understanding of the manner in which the carrier-carrier (e-e, h-h, e-h) interactions influence the cooling and response to externally applied electric fields of the photoexcited carriers at different excitation energies and levels.

In this paper, we examine the energy dependence of the role of e-e, h-h, e-h, and carrier-phonon interactions on the cooling rates of electrons and holes. The EMC approach is used to avoid making any assumptions on the form of the distribution functions for the holes and electrons which has been a standard practice in previous theoretical attempts to investigate the cooling process[8,9]. With its built-in time evolving distribution, the EMC approach makes it possible to examine the details of the cooling process on picosecond and sub-picosecond time scales. In an earlier study, we examined the concentration dependence of an e-h plasma excited by a 200 femtosecond, 1.64 ev laser pulse[10]. It was found that the e-h interaction provided a significant channel through which the electron energy is transferred to the lattice particularly at high excitation levels where the electron-phonon interaction is strongly screened and the e-h scattering events are more frequent. The cooling rates were lower than those predicted by a model that assumed a parameterized distribution function with equal electron and hole temperatures and includes the hot phonon effects. More, recently we also examined the hot phonon and e-h interaction using EMC and showed that for times less than 2 ps the e-h interaction controlled the relaxation process while for times greater than 2 ps the not phonons slowed the cooling rates[11].

In the present study, situations where the e-h plasma is photoexcited with excess energies of 40, 200, and 300 meV are investigated. These energies correspond to situations where the electron energy (E_e), the hole energy (E_h), and the LO phonon energy (E_0) satisfy the following conditions:

(1) $E_e < E_0$ and $E_h < E_0$,
(2) $E_e > E_0$ and $E_h < E_0$,
(3) $E_e > E_0$ and $E_h > E_0$,

respectively. Hot phonon effects are not considered. The density of the e-h plasma is assumed to be $5 \times 10^{16} cm^{-3}$, and the finite duration of the pulse is ignored. From the results of this study, conclusions are drawn about the influence of the initial energy distribution of the photoexcited carriers on the manner in which c-c interactions achieve thermalization of the carrier distributions.

Theory

The EMC proceedure used in this study is the same used in investigating the role of e-h interaction[11,12]. Thus only its main features will be summarized for the sake of completeness. The band model consists of three nonparabolic valleys for the conduction band and a parabolic heavy hole valence band. The light hole band is ignored, partly because of its small density of states and to reduce the computation time. Figure 1, shows the flow chart of the EMC program used in this investigation. The first step corresponds to specifying the material parameters, initial energy distribution of the carriers, pulse shape and duration, room temperature, value of the electric field, and which c-c interactions are active. The scattering rates for acoustic phonons, deformation potential, intervalley and intravalley scattering processes for electrons and holes are then calculated and tabulated. On the other hand, the scattering rates of c-c and screened carrier-phonon interactions, which depend on the evolving distribution function of the carriers are calculated every 50 femtosecond initially and at longer time intervals as the time proceeds. For finite pulse duration, the number of the carriers is adjusted according to the pulse shape with an initial energy determined from the energy of the pulse taking into account the nonparabolicity of conduction band. The dynamics of the electrons and holes are then examined using the regular EMC approach. Coupling between the electron EMC and the hole EMC programs is accomplished through the e-h interaction which depends on the energy and momentum distribution of the electrons and holes. The average energy of the electrons and holes, energy-loss-rates through e-h and electron-phonon interactions are then calculated at the end of each iteration or at specified time intervals.

The electrons and holes are assumed to interact among themselves and with each other through an exponentially screened Coulomb potential. The screening length is calculated self consistently in the static and long wavelength limit of the random phase approximation. Unlike previous models for e-h scattering, no a priori assumptions are made on the magnitude of the energy exchange between the electrons and holes in an e-h scattering event. The expressions for the scattering rates for c-c and screened carrier-phonon interactions are given in reference 12. Figure 2, shows the scattering rates for e-e, e-h, h-h interaction processes, assuming a Maxwellian distribution for the electrons and holes at temperatures of 1000 K and 160 K respectively as a function o. the average wavevector kz. From this figure, it is clear that the h-h scattering rates are much stronger than the e-e rates due to the larger density of states in the heavy hole band. Also notice that the interaction of an electron with a hole plasma is stronger than the rate at which a hole interacts with a sea of electrons[13]. This is related to the large density of final states for holes compared to that of electrons. The screening of the carrier-phonon interactions reduces the scattering rates significantly, especially at high carrier concentrations and low temperatures. Additionally, the presence of the relatively cold hole plasma, which is the case in a photoexcitation process, leads to stronger screening even when the energy of the electrons is sufficiently high[12]. In the present analysis the acoustic phonon scattering is assumed to be elastic which is a good approximation at lattice temperature of 77 K. However, for excitation by total excess energy of 40 meV, the use of an inelastic acoustic scattering model is more appropriate because the relaxation through the optical phonon emission is severely diminished. Furthermore, the interaction between the L-valley electrons and holes is neglected, because the population of the electrons the upper valleys is negligible. The elastic impurity scattering is ignored in our calculation, since we assume that all of the carriers are generated by the photoexcitation process. In order to examine the roles of e-e , h-h, and e-h interactions more closely, the phonon system was assumed to be in equilibrium.

Results and discussion

The cooling rates of carriers photoexcited with total excess energies of 40, 200 and 300 meV, corresponding to photon energies of 1.55, 1.71, and 1.81 eV respectively, were examined under four different combinations of active scattering processes. First the simulation was carried out assuming only optical phonon interactions were active. Curve 1 in Figures 3 through 5 corresponds to this situation. Notice that curve 1 in Figures 3 and 4b, which represents cooling of electrons and holes with initial energies below the phonon emission threshold E_0 (36.4 meV), does not change with time since no phonon emission is possible and at 77 K and the probability for optical phonon absorption is small. However, when the initial energy of the electrons is above E_0 (Figures 2a and 3a), the electrons emit integral multiple of LO phonons and end up with energies equal to 27.9 meV and 4.25 meV at 4 picosecond after excitation by 1.71 and 1.81 eV laser pulses, respectively. Similarly the holes reach a steady state energy of 5.8 meV which corresponds to emitting a single optical phonon (see Figure 3b). After reaching these energies no further cooling is possible to reach equilibrium with the lattice, since the acoustic scattering is assumed to be elastic.

The situation where the e-e and h-h interactions were active, in addition to the optical and acoustic phonons, was examined next. Curve 2 in these figures shows the cooling process for the electrons and holes. The most interesting phenomena in Fig. 1 is that, the e-e interaction enhances the cooling of the electron system from an initial energy of 35 meV, which is just below the phonon emission threshold, to a steady state energy 13.5 meV that is slightly above the room temperature. This enhancement occurs because some electrons gain enough energy from interacting with other electrons and then cool down by emitting LO phonons. On the other hand, the h-h interaction coupled with optical phonon absorption drives the hole system to thermal equilibrium with the lattice. The average energy per hole changes slowly from an initial energy of 5 meV to a final energy 8.7 meV which is just below E_{th}. In this case each hole gains an average energy of 3.7 meV from the lattice which means that energy of a single optical phonon is shared by about 10 holes. This is possible because the strength of the h-h interaction spreads the distribution of the holes quickly in energy, so that when a hole absorbs an optical phonon, it is more likely to share the gained energy with an other hole than emitting an optical phonon. This results in two holes whose energies are below E_0, but are relatively hotter than the rest of the holes, so that through further h-h scatterings the extra energy will be distributed among the holes. When the electrons are excited by 1.71 and 1.81 eV laser pulses, the e-e interaction leads to a slower cooling rate. This happens because the interaction among the electrons quickly redistributes the electrons into high and low energy regions. Consequently, many electrons end up with energies below the phonon emission threshold, while those which end up at higher energy regions essentially have the same probability for emission of optical phonons. Those electrons which have energies below the phonon emission threshold can cool down only by giving up some of their energies to other electrons or gaining enough energy to emit an optical phonon. Comparing curve 2 in Figures 4a and 5a, we see that the manner in which e-e interaction influences the cooling rate is also dependent on whether the final energy after emitting the LO phonons (assuming no c-c interactions) is below or above E_{th}. When the e-h plasma is excited by excess energy of 200 meV, the electrons would reach a steady state energy of 28 meV by emitting LO phonons only, which is not in equilibrium with the lattice. Consequently the electrons exchange energy among themselves and emit LO phonons in a manner similar to the situation in Figure 3a. This explains why the cooling curve 2 in Figure 4a drops below curve 1 beyond 2 ps. On the other hand curve 2 always remains above curve 1 in Figure 5a, and the electron ensemble reaches a steady state energy of 19.4 meV which is above E_{th}. The cooling curve for the holes in Figure 4a behaves in the same way the the electrons do in Figure 3a, because in both cases the initial energies of the carriers are less than E_0 and greater than E_{th}. When the initial energy of the holes is above E_0, the energy exchange among the holes leads to faster equilibrium with the lattice as can be seen in Figure 5b (curve 2). The holes retain some part of the initial energy instead of losing it to the lattice via optical phonon emission.

The case where only electron-phonon and e-h interactions were active was then investigated. From examining curve 3 in Figure 3b, we see that the average energy of the holes increases rapidly during the first picosecond from 5 meV to 12.5 meV as a result of the energy transferred from the electrons. Also the cooling curve for curve 3 in Figure 3a is below that of curve 2 reflecting the fact that e-h interaction tends gradually shift the electron population to lower energy states by transferring the energy to the holes. The cooling rates exhibit the same trend for all the cases examined as can be seen from the relative positions of curves 2 and 3 in Figures 3 through 5. Finally the more realistic situation where all the scattering mechanisms e-e, e-h, h-h, and carrier-phonon are present was investigated. From curve 4 in Figure 3a it is obvious that the coupling of e-e and e-h interactions leads to a faster cooling of the electrons to room temperature. The same trend is true for the holes, so that the

energy of the holes is intermediate between the situations where h-h or e-h interaction is active only. However, it is clear form examining curves 3 and 4 in Figures 3 through 5, that the cooling rates are primarily determined by the e-h interactions, since the cooling curve in this case closely follows the situations where only e-h and carrier-phonon interactions were active.

Conclusion

The c-c interactions are sensitive to the initial energy distribution of the carriers and they influence the carrier-phonon interactions so as to achieve a faster thermalization of the carrier distribution. when the initial energy of the carrier is below the phonon emission threshold c-c interactions stimulate the optical phonon emission or absorption process depending on whether the initial energy of the carrier is above or below the thermal energy respectively. The e-h interaction plays a dominant role in the relaxation process even when excitation energy is below the LO phonon emission threshold.

Acknowledgment

References

1. R. L. Fork, Physica 134B, 381-388 (1985).
2. R. A. Hopfel, J. Shah, P. A. Wolf, and A. C. Gossard, Phys. Rev. Lett. 56, 2736 (1986).
3. M. R. Junnamarker and R. R. Alfano, Phys. Rev. B 34, 7045-7062 (1986).
4. J. Degani, R. F. Leheny, R. E. Nahory, and J. P. Heritage, Appl. Phys. Lett. 39, 569 (1981).
5. D. D. Tang, F. F. Fang, M. Scheuermann, and T. C. Chen, Appl. Phys. Lett. 49, 1540 (1986).
6. R. A. Hopfel, J. Shah, and A. C. Gossard, Phys. Rev. Lett. 56, 765 (1986).
7. J. C. Tsang and J. A. Kash, Phys. Rev. B 34, 6003 (1986).
8. W. Potz and P. Kocevar, Phys. Rev. B 28, 7040 (1980).
9. M. Asche and O. G. Sarbei, Phys. Stat. Sol. (b) 126, 607 (1984).
10. M. A. Osman, U. Ravaioli, R. Joshi, W. Potz, and D. K. Ferry, in Proc. 18th Intern. Conf. Phys. Semiconductors, Stockholm, 1986, Ed. by O. Engstrom, in press.
11. M. A. Osman, H. L. Grubin, P. Lugli, M. J. Kann, and D. K. Ferry, to be published in the proceedings of the Second Topical Meeting on Picosecond Electronics and Optoelectronics, Lake Taho, 1987.
12. M. A. Osman, Ph.D. dissertation, Arizona State University, Tempe, August 1986 (unpublished).
13. M. A. Osman and D. K. Ferry to be published.
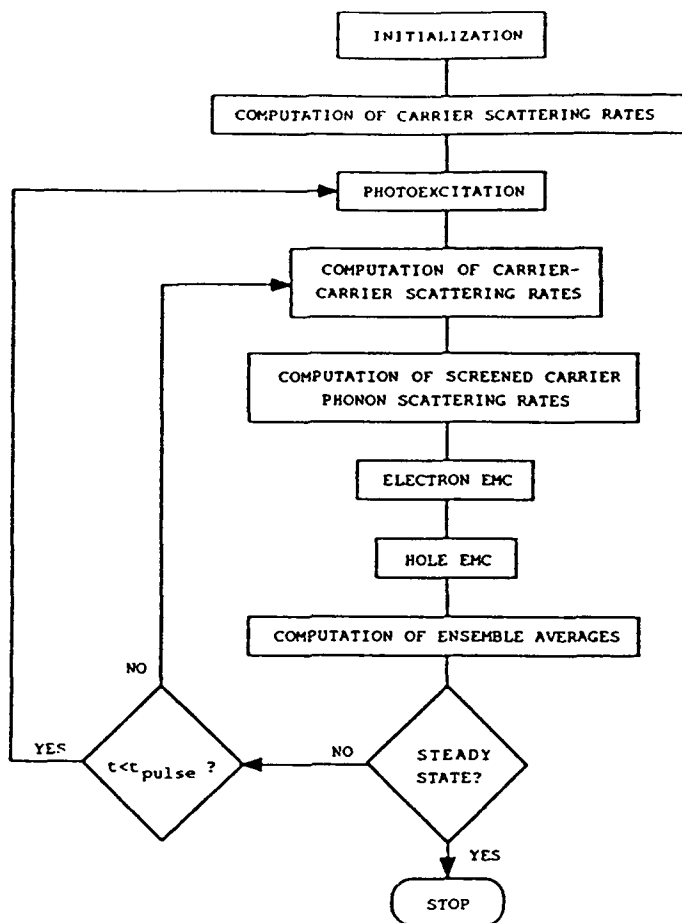
Figure 1. Flow chart of the EMC program.
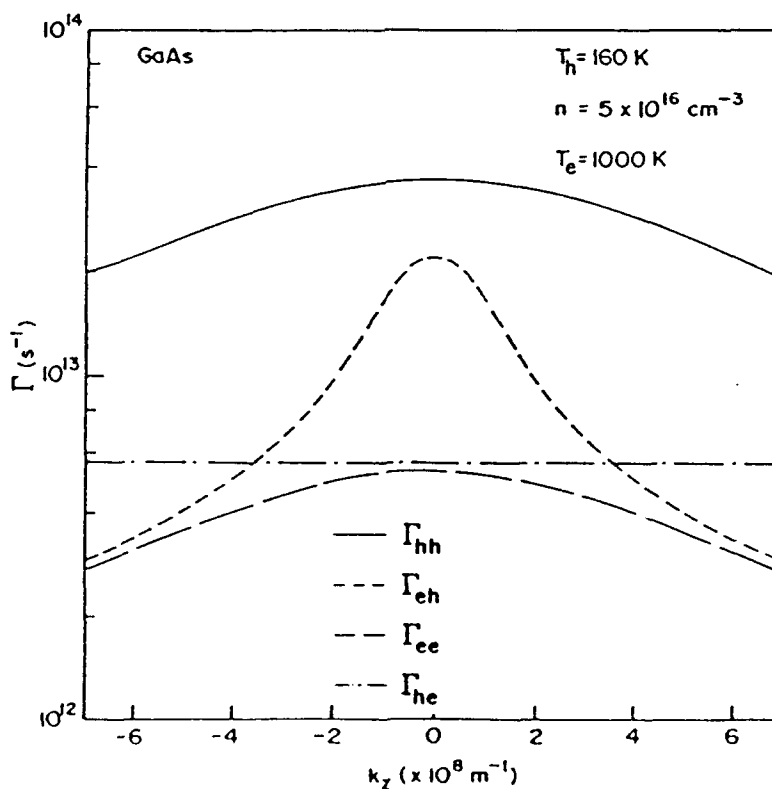
Figure 2. Carrier-Carrier scattering rates as a function of the average kz.
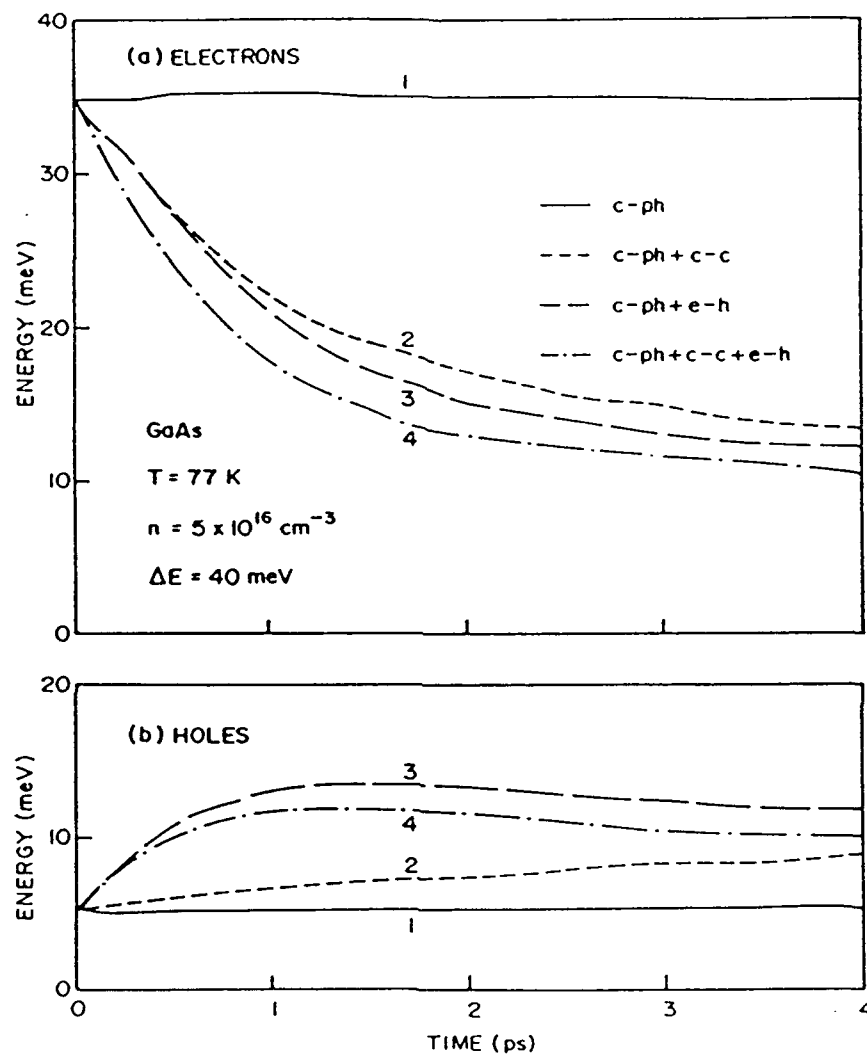
Figure 3. Time evolution of the carrier energy for different active scattering mechanisms (a) electrons, (b) holes.
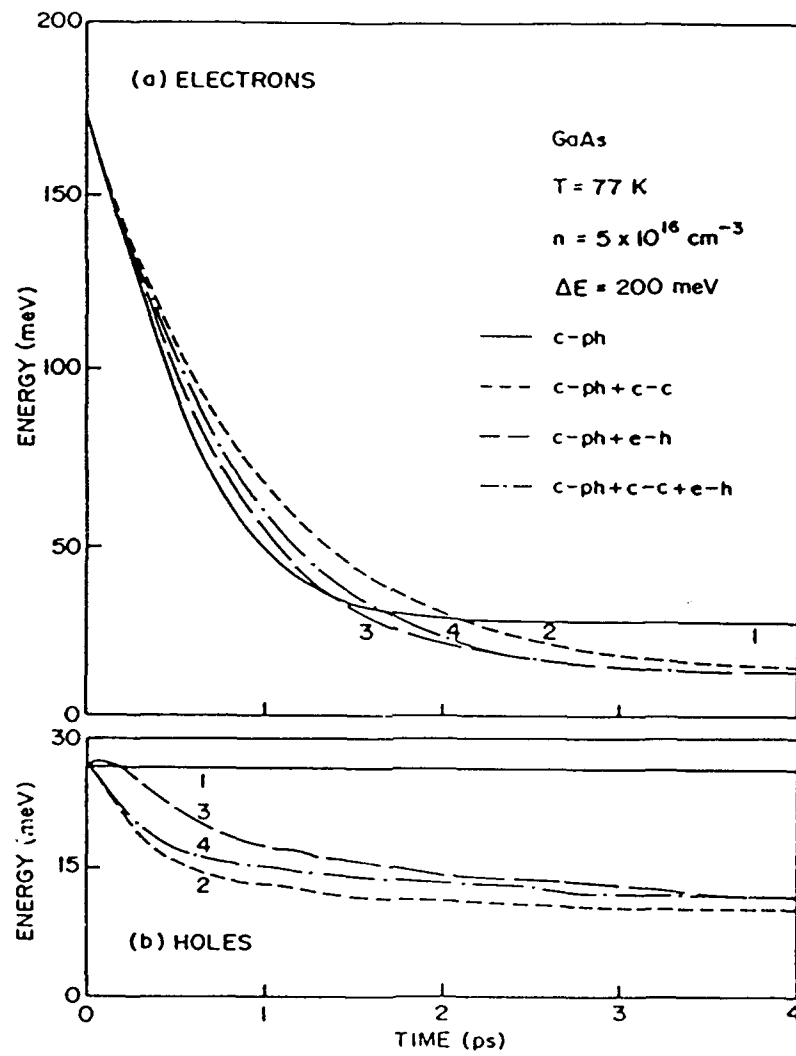
Figure 4. Time evolution of the carrier energy for different active scattering mechanisms (a) electrons, (b) holes.
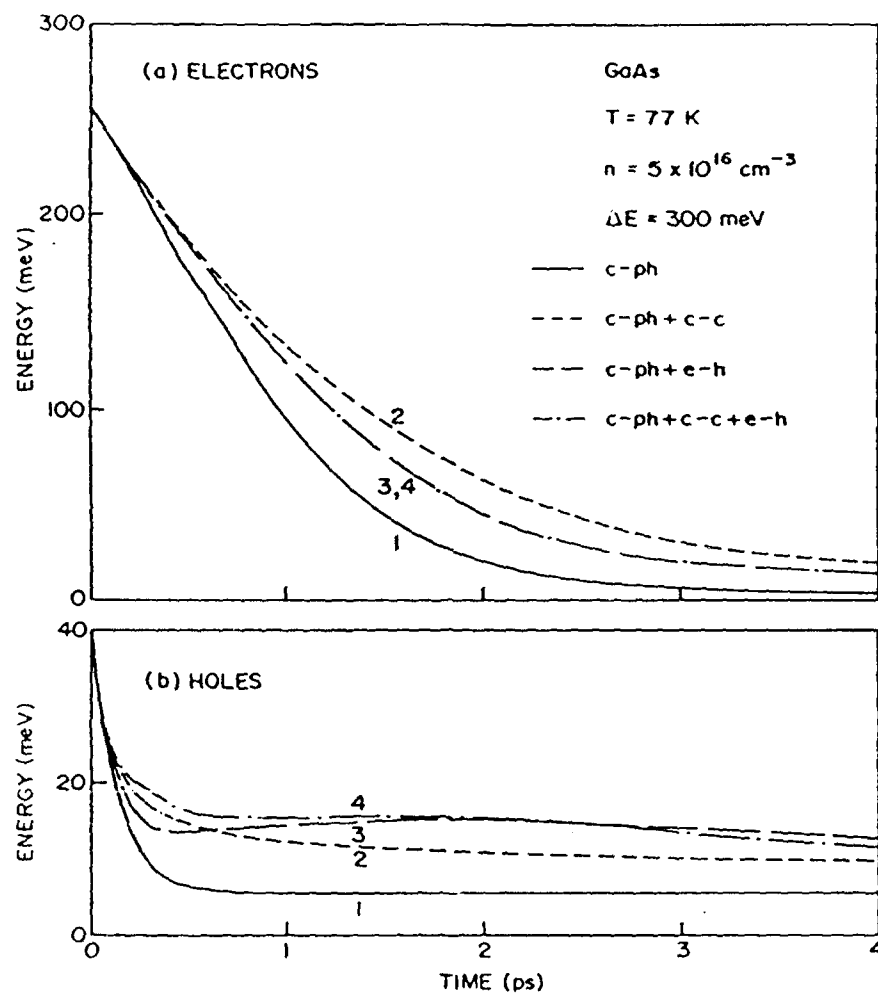
Figure 5. Time evolution of the carrier energy for different active scattering mechanisms (a) electrons, (b) holes.

Monte Carlo Investigation of Hot Photoexcited Electron Relaxation in GaAs*

M. A. Osman and H.L. Grubin
Scientific Research Associates, Inc.
Box 1058, Glastonbury, CT 06033

P. Lugli
Dipartimento di Fisica e Centro Interuniversitario di Struttura della Materia della Universita, Modena, Italy

M.J. Kann and D.K. Ferry
Center for Solid State Electronics Research,
Arizona State University, Tempe, AZ 85281

Recently, several authors [1-3] have investigated the dynamics of electron-hole plasma generated by picosecond and sub-picosecond laser pulses in GaAs. These measurements have indicated the importance of electron-hole interaction in the cooling process (especially in the presence of a cold hole plasma) and the existance of non-equilibrium phonon distributions as a result of cooling of photoexcited electrons and holes. Previous theoretical studies of the cooling process have ignored the presence of the upper valleys even when the energy of the excited electrons exceeds the energy separation between the central and the upper valley. Additionally, assumptions had to be imposed on the form of the distribution function of the electrons and phonons. Ensemble Monte Carlo (EMC) techniques avoid these assumptions and have been used to investigate the effects of electron-hole (e-h) interaction [4] and non-equilibrium phonons [5] on the relaxation rate of photoexcited electrons. In this study we present the first EMC calculation to account for both the e-h interaction and hot phonon effect.

The Monte Carlo model takes into account the e-h , carrier-phonon interactions and the non-equilibrium LO phonon-electron interactions. The band model consists of three nonparabolic valleys for the conduction band and a parabolic heavy-hole valence band presence. The parameters used in the EMC calculations are the same as those used in references [4] and [5]. Electron-electron, electron-plasmon and the screening of electron-phonon interactions have been negelected. Initially, to avoid the complications arising from the transfer of electrons to the upper valleys, we have examined the roles of hot phonons and electron-hole interaction for excitation energy of 1.8 eV. Under this situation, the energy of the excited electrons is 250 meV while that of the holes is 40 meV, so that the change in the LO phonon population is mainly due the emission of LO phonons by the electrons. The time evolution of the LO-phonon distribution $N(q)$ is calculated as a function of the wavevector q from the Monte Carlo simulation using the approach outlined in [5]. Instead of recalculating the scattering rates for the LO-phonons using the perturbed phonon distribution, we have used a Large value for $N(q)$ and used the rejection technique to accept the final state after each scattering event involving an LO phonon.

The relaxation of electrons excited by a 1.8 eV laser pulse was examined for an excitation level of 5×10^{16} cm^{-3} to understand the manner in which e-h interaction and hot phonons influence the cooling process. From Fig. 1, it is clear that when the e-h interaction is ignored, the hot phonon effect is significant only for times longer than 1 ps. Curve 1 in Fig. 2 shows that between 0.6 and 2 ps the energy-loss-rate through the emission of LO-phonons is reduced when the phonon heating is taken into account compared to the situation where hot phonons are negelcted (curve 1). When the e-h interaction is taken into account, the cooling rate slows for times less than 2 ps as shown by curve 3 in Fig. 1. When both the e-h interaction and hot phonons are considered (curve 4), the cooling rate is slightly faster for times below 1 ps partly due the initial overestimation of the enhancement of the phonon emission process inherent in using a single large value for N(q). However, it is obvious that the cooling process for times larger than 2 ps is controlled by the hot phonons, which lead to a slower cooling rate. Additionally, no significant change in the energy of the photoexcited holes was noticed, when the electron-hot-phonon interaction was considered. The situation where hole-hot-phonon interaction is taken into account might lead to an increase in the hole energy as a result of the phonon reabsorption process. This work is underway and the results will be presented elsewhere.
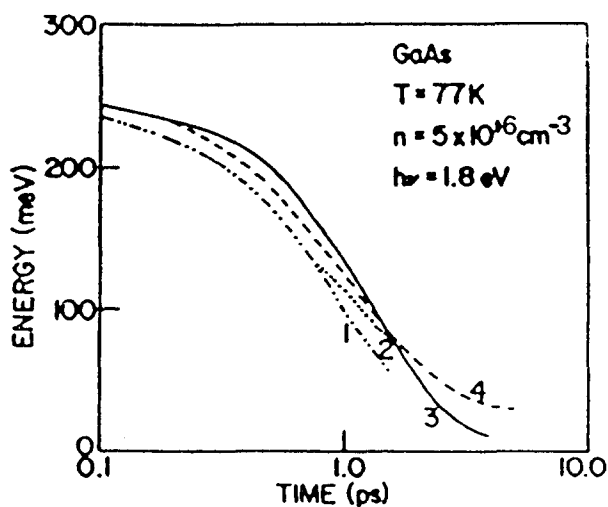
Fig. 1. Time evolution of the electron energy: 1) e-ph, 2) e-ph (hot), 3) e-ph + e-h, 4) e-ph (hot) + e-h.
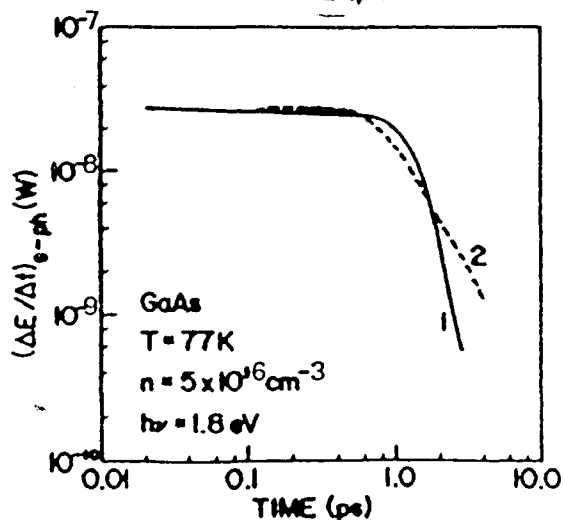
Fig. 2. Time dependence of the electron-energy-loss rate: 1) w/o hot phonons, 2) with hot phonons.

The situation where the electrons were transferred to the upper valleys, was investigated for the case of excitation by a 588 nm laser pulse using EMC and taking into account the carrier-phonon interactions only. The fraction of electrons in the central valley is shown in Fig. 3 as a function of the time after the application of the pulse. From Fig. 3, it is obvious that the simple picture of electrons cascading down the central valley by emitting LO-phonons is not complete. The electrons rather transfer to the upper valleys and cool down in these valleys by emitting phonons through polar-optical and deformation potential interactions as shown in Fig. 4 (curves 2 and 4). The energy loss rate through LO phonon emission is dominated by the central valley electrons for times greater than 1 ps and reach a maximum after 2 ps. Note that after 2 ps about 80% of the electrons are in the central valley. The energy of the electrons in the central valley (curve 1 in Fig. 5) remains above 100 meV up to 3 ps after the excitation and an increase in the average energy occurs at 2 ps due to the return of the electrons to the central valley. The initial rapid drop in the average kinetic energy of the electron ensemble (curve 2 in Fig. 5) is due to the transfer of a large fraction of the electrons to the upper valleys as can be seen in Fig. 3. Curve 3 in Fig. 5 indicates that the cooling of the electron in the L-valley is not negligible.
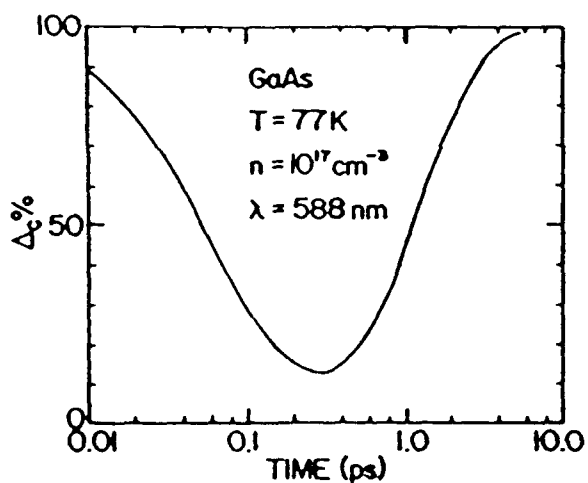
Fig. 3. Time evolution of the fraction of electrons in the central valley.
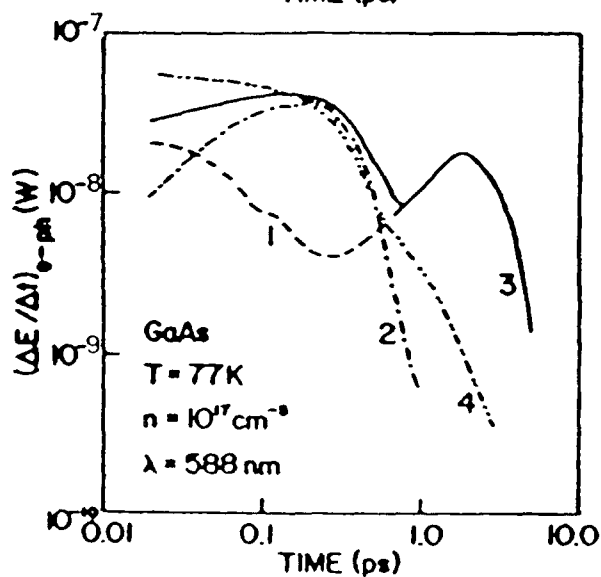
Fig. 4. Time dependence of the electron-energy-loss rate:
1) LO (Γ), 2) LO (L),3) total LO
4) Intervalley scattering
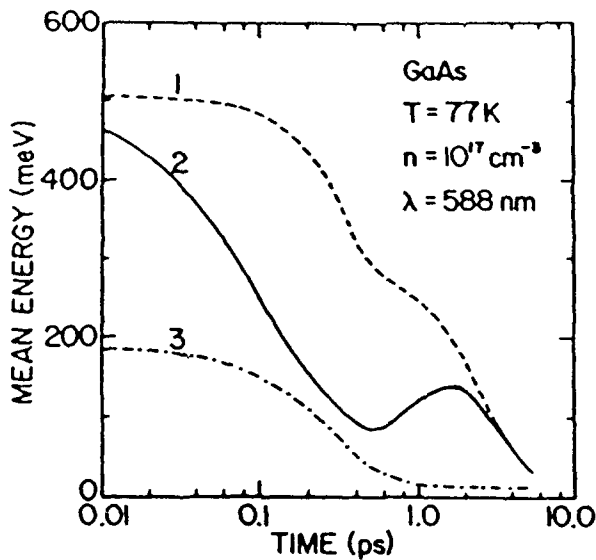
Fig. 5. Mean kinetic energy of the electrons in: 1) central valley, 2) central + L-valley, 3) L-valley.

In conclusion, have shown that the cooling process of the photoexcited electrons is controlled by e-h interaction for times less than 1 ps and by hot phonons for times greater by 2 ps. In addition, we have demonstrated that the cooling of the electrons in the upper valleys are important for high excitation energies.

References

1. J. A. Kash, J. C. Tsang, and J. M. Hvam, Phys. Rev. Lett. 54, 2151 (1985).
2. J. Shah, A. Pinczuk, A. C. Gossard, and W. Weigmann, Phys. Rev. Lett. 54, 2045 (1985).
3. C. L. Collins and P. Y. Yu, Phys. Rev. B30, 4501 (1984).
4. M. A. Osman, U. Ravaioli, R. Joshi, W. Potz, and D. K. Ferry, in Proc. 18th Intern. Conf. Phys. Semiconductors, Stockholm, 1986, Ed. by O. Engstrom, in press.
5. P. Lugli, C. Jacoboni, L. Reggiani, and P. Kocevar, submitted for publication.

TRANSFER OF COPYRIGHT TO THE SOCIETY OF PHOTO-OPTICAL INSTRUMENTATION ENGINEERS

Title of Paper:

SPIE CODE NO.

Ultra transients in semiconductors

793 10

Author(s):

H.L. Grubin, M.A. Osman and J.P. Kreskovsky

(signature)	_(signature)_	
Signature	Signature	Signature
Harold L. Grubin	Harold L. Grubin	
Print Name	Print Name	Print Name
Vice President	Vice President	
Title Date	Title Date	Title Date

Signature	Signature	Signature
Print Name	Print Name	Print Name
Title Date	Title Date	Title Date

>M/7-86

Ultrashort transients in semiconductors

H.L. Grubin, M.A. Osman and J.P.Kreskovsky
Scientific Research Associates, Inc.
P.O. Box 1058, Glastonbury, Connecticut 06033-6058

Abstract

Over the past decade a variety of physical phenomena have emerged which are exercising profound constraints on the speed of devices whose primary transient is electronic in origin. For example, the broad high frequency extrapolations of high speed arising from velocity overshoot were found to be limited by the effects of space charge and the imposition of fields that take a finite time to reach their specified level. Much of the latter problem may be minimized through optical processes, but in its place rests the constraints of electron-hole interaction and other processes that may require quantum coherence for implementation. In the case of electron-hole interaction it has been found that it provides an important channel through which electron energy is transferred to the lattice, and that the electron cooling rates at high hole concentrations are higher when the electron-hole interaction is included, than when it is ignored. At low hole concentrations the reverse is true. The effect of electron-hole interaction also influences velocity overshoot. In the case of quantum coherence, of interest is the effect, e.g., of temporal scatterers (phonons) on the time dependent wave functions and modifications of interference effects. The influence of space charge on transient velocity overshoot, the effects of electron-hole interaction on the relaxation of an electron-hole plasma, and the effects of time dependent and time independent scattering on quantum interference will be discussed.

Introduction

The early suggestions of Ruch[1] and subsequent calculations of others, have suggested that high mobility semiconductors were capable of extraordinary high speeds (10^8cm/sec) provided they were sampled for sufficiently short time intervals. For submicron and ultrasubmicron devices, this translates into time scales of 0.1 ps, and cutoff frequencies of 10^3GHz. While these predictions are clearly impressive, of serious concern is an inability of realizing the modest goal of room temperature operation at 100GHz, and speeds of less than 5ps, within the framework of a three terminal device. The difficulty of achieving anything close to that offered by the Ruch calculations has been discussed by a number of workers over the past several years. Perhaps the simplest statement summarizing these studies is that Ruch's assumptions are not realized experimentally. These assumptions include: 1) uniform fields within the device, 2) zero rise time of the electric field, and 3) neglect of displacement currents. Because of the importance of these assumptions, each will be briefly reviewed, along with the influence of nonequilibrium holes and phonons. Additionally the influence of quantum mechanical constraints on transport will be discussed.

Nonuniform fields

Electric fields in devices are nonuniform, as illustrated for the three terminal permeable base transistor (PBT). The PBT as currently configured, is a micron-length structure with ultra-submicron features. A section of the PBT is shown in figure 1, along with its an equivalent planar FET. As shown in figure 2[2], for a bias of one volt between the emitter and the collector the potential distribution indicates the expected field non-uniformity. For example, near the cathode, the field is low, increasing to a value near 20 kv/cm and finally approaching a low value near the collector contact. In uniform field studies the predicted peak velocity approaches 10^8cm/sec, followed by a decay to steady state. The peak velocity is achieved with a carrier concentration that is close to its equilibrium value. Decay is dominated by electron transfer, with a weaker component due to momentum relaxation. In steady state the gamma valley carrier velocity is high, while the concentration of gamma valley carriers is low. In the PBT the initial high peak in current associated with velocity overshoot is not achieved. Rather the high velocity is often achieved at the expense of carrier concentration. This is displayed in figures 3 and 4 which show the distributon of carrier density and velocity for the PBT. It is worthwhile noting that while the distribution of charge and velocity are not as encouraging as that from the uniform field calculations, the current voltage levels are between a factor of 3 and 4 above that obtained from the drift and diffusion equation calculations.

Finite rise time contributions

A second issue is the zero rise time. In virtually all transient transport calculations, overshoot phenomena is obtained by examining the response of carriers to an electric field that is suddenly turned on. Under conditions of finite rise time, relaxation effects are coupled to the increasing field and the peak velocity is diminished, as shown in figure 5[3]. We note that there is a qualitative similarity between the finite rise time and the response of carriers to a gradual spatially-varying electric field.

Displacement current effects

To carry the discussion of non-uniform fields even further, calculations have been performed for the paradigm two-terminal $N^+N^-N^+$ submicron structure shown in figure 6[4]. The structure results in non-uniform fields and when transients are computed it is a current transient that is involved rather than a velocity transient. Indeed, in virtually all structures it is a current transient rather than a velocity transient that is measured. As shown in figure 7, there is a peak value to the current that occurs within the time scale over which velocity overshoot may be expected. However, the details of the transient show that the initial time dependent behavior is associated entirely with field rearrangement, and that the transient is entirely a displacement current transient[4].

Nonequilibrium electron hole phonon interaction

In overshoot studies most inelastic collisions are treated assuming the emission of a single phonon and that the emitted phonon is in equilibrium with the lattice. For energetic electrons, a sequence of as many as ten to fifteen phonons maybe emitted as the electron is relaxing, and as a result an excess number of phonons may result in a lack of phonon equilibrium with the lattice. This phenomena is illustrated for an unbiased slab of gallium arsenide subject to a uniform photoexcited laser pulse of energy 1.8 ev. Under this situation holes as well as electrons are excited. The energy of the excited electrons is 250 Mev, while that of the holes is 40 Mev. The time evolution of the resulting electron, hole and phonon distribution is calculated using Monte Carlo techniques. Figure 8 shows the relaxation of the electrons to equilibrium with and without hot phonons. As seen, the inclusion of the hot phonons results in a delay of the relaxation to steady state.

The situation in the presence of finite fields and equilbrium phonons is discussed next and shown in figures 9 and 10[5]. These calculations were performed assuming nonequilibrium interactions between the gamma valley electrons and holes. L-valley hole interaction was not examined. The first point of note is that the electron hole interaction is an inelastic collision. In addition, since both electrons and holes are responding to the presence of the applied fields, momentum scattering of the electrons and holes can result in a reduction of the electron velocity, and in some cases turn the electron in the "wrong" direction. The net result is a decrease in the peak electron velocity. It is important to note that while the peak velocity is lower for the calculation with the electron-hole interaction, there is a cross-over where, for a certain interval of time, the carrier velocity in the presence of holes exceeds that obtained in the absence of holes. This arises because the electron hole interaction is an efficient method of energy exchange and as a result, the electrons are not energetic enough to transfer into the L-valley. In other words, there is a greater fraction of electrons that are retained in the gamma valley when holes are included than when they are not included. The possibility exists that the velocity of injected electrons in p-type material may, for a range of values of electric field, be higher than that of thermally generated carriers in N-type material.

Quantum contributions

It will be noticed that all of our discussion has been confined for electric fields and scalar potentials. It has been presumed that we are dealing with a gauge invariant system, and that the choice of gauge is not relevant. As discussed by Aharanov and Bohm and[6], the vector potential can lead to observables such as interference patterns for electrons that traverse ostensible field-free regions. In other words, interference patterns are accounted for in terms of differences in the vector potentials encountered by electrons on two opposite sides of a flux line. To see how these effects enter into the problem it is useful to re-examine an approach to quantum transport introduced earlier[7].

In reference 7, which introduced the Wigner function as an approach to quantum transport in electron devices, a set of hydrodynamic balance equations was introduced. While the hydrodynamic balance equations were developed for a system of particles, the equations may be reduced to examine transport for carriers not subject to scattering. Thus, the single particle wave functions are of interest.

Determining the single particle wave functions requires solutions to the single particle Schrodinger's equation. Rewriting Schrodinger's equation in terms of a hydrodynamic model requires solutions to the following two equations:

$$\frac{\partial \rho}{\partial t} + \text{div} J = 0 \tag{1}$$

$$m\frac{d\upsilon}{dt} = -\nabla(Q+V) \tag{2}$$

where $V(x)$ is a classical scaler potential and Q is the quantum potential

$$Q = -\frac{h^2}{2m} \frac{\nabla^2 \sqrt{\rho}}{\sqrt{\rho}} \tag{3}$$

The reference 7 study suggests that both $\rho = \Psi^*\Psi$ and $J = \rho\upsilon$ be regarded as the basic physical quantities representing properties of the continuous fluid. It is noted that the vector potential has dropped out of the basic equations, which have a completely classical form.

Philippidis, Bohm and Kaye (PBK)[8] argued that the above equations were incomplete. The most important objection arose because certain subsidiary conditions were not satisfied. To see this, the wave function for the system is written as $\Psi = R\exp[iS/h]$, with the carrier velocity expressed as

$$\upsilon = \frac{\nabla S}{m} - \frac{e}{mc} A \tag{4}$$

If it is now required that the wave function be single valued, then we are forced to admit the condition

$$\oint \nabla S \cdot dx = \oint \left(m\upsilon + \frac{e}{c} A\right) \cdot dx = nh \tag{5}$$

This means that the observable movements of electron transport in a device must necessarily be restricted in such a way that the integral of its velocity around a circuit containing a flux line depends upon the flux within the circuit. AB argue that there is no intuitively clear justification for the equation (8) constraint. Rather, they propose to assume that the vector potential has a physical meaning.

Figure 11 is a plot of the quantum potential for an electron passing through a dual slit in the absence of a magnetic field. The smooth parabolic hills are in the immediate vicinity of the slit. Particles subject to such a potential accelerate smoothly in the forward and sideward direction. This can be seen from the family of velocity trajectories displayed in figure 12. Note, near the slits where the wave functions are localized packets, the quantum potential exhibits two broad peaks. As the wave packets recede and spread, they begin to overlap significantly and their interference properties come into play. The shape of the quantum potential, a long way from the slits, is determined by the boundary conditions, the shape and size of the slits. It is through the force exerted by the quantum potential on the particle that the boundary conditions are made physically significant far from the slits. For the case when a flux line is place between the slits, the quantum potential is altered, as shown in figure 13. The velocity trajectories are altered as shown in figure 14. The important point to notice is that the trajectories are displaced.

The above discussion reached initial device implementation by Datta, et al.[9], who conceptually altered the path length of a beam of particles as shown in figure 15, where the slit in the GaAs between the AlGaAs is exaggerated. Additionally they altered the phase through application of a magnetic field. These alterations lead to interference in the downstream carrier wave functions. This result, as discussed in reference 9, is the appearance of oscillations in the conductance of the system.

An important point not addressed in reference 9 is the temporal duration of the conductance oscillations. It must be considered that the sporadic appearance of phonons will result in a loss of coherence and limit constructive interference to time scales shorter than 100 femtoseconds. For time scales longer than 100 femtoseconds, the constructive interference effects will certainly be diminished. Whether they will disappear is a matter to be determined. Here we note that coherence is essential for quantum well devices. There are other issues that may tend to mitigate the appearance of the AB effect such as elastic scattering arising from surface roughness.

Conclusions

The driving force toward high speed devices is limited by the presence of carrier-carrier scattering, phonons and submicron feature sizes that tend to reduce the coherence of scattered waves. While these effects do not lessen the importance of the nonequilibrium transients, unless these effects are incorporated into the design of the device, applications may be limited.

Acknowledgements

References

1. J.G. Ruch, IEEE Trans. Electron Devices ED-19, 652 (1982).

2. J.P. Kreskovsky and H.L. Grubin, Proceedings of NUMOS I (1987).

3. H.L. Grubin, J.P. Kreskovsky, Surface Sci. 137, 594 (1983).

4. H.L. Grubin, J.P. Kreskovsky, VLSI Electronics Microstructure Sci. 10, 237 (1985).

5. M.A. Osman, PhD Thesis, Arizona State University (1986).

6. Y. Aharonov and D. Bohm, Phys. Rev., 115, 485 (1985).

7. G.J. Iafrate, H.L. Grubin, and D.K. Ferry, J. de Physique, C7, 307 (1981).

8. C. Philippidis, D. Bohm, and R.D. Kaye, Il Nuovo Cimento, 71B, 75 (1982).

9. S. Datta, M.R. Mellcoh, S. Bandyopadhyay, and M.S. Lundstrom, Appl. Phys. Letts. 48, 487 (1986).
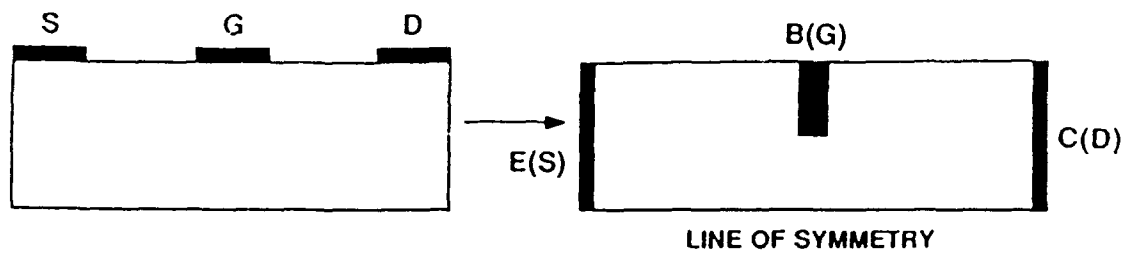
Figure 1. Schematic of FET and PBT.

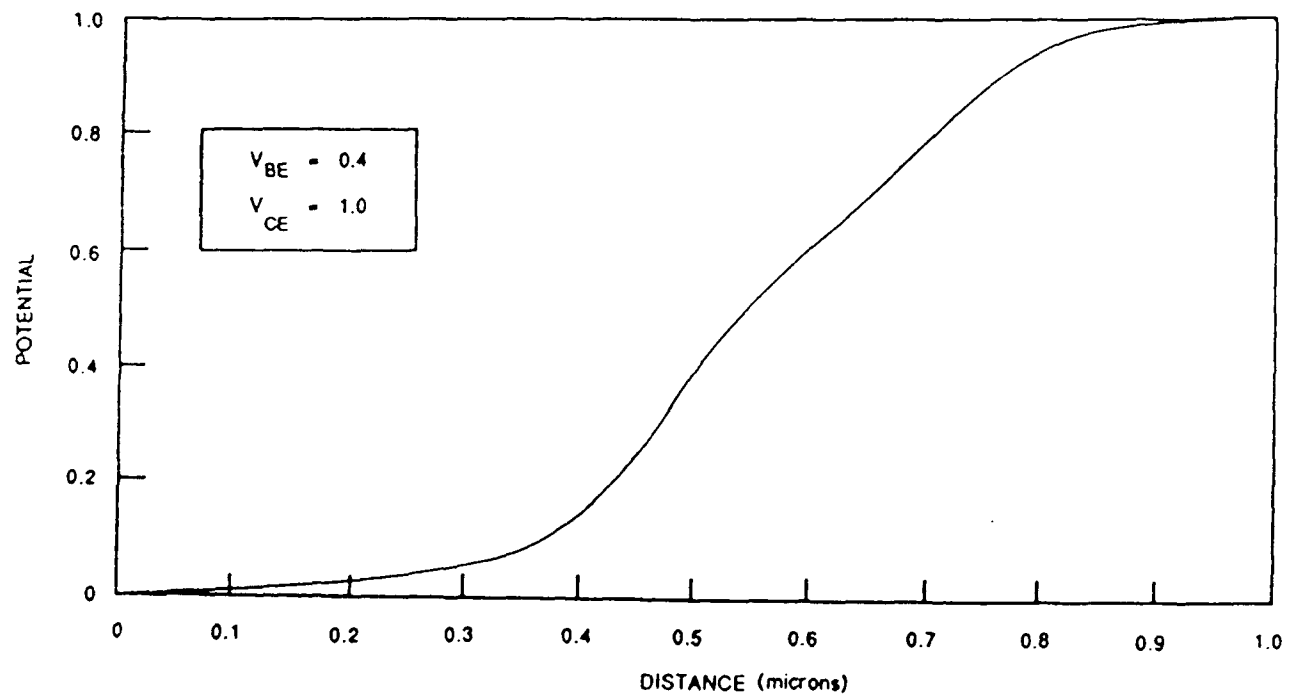$V_{BE} = 0.4$

$V_{CE} = 1.0$

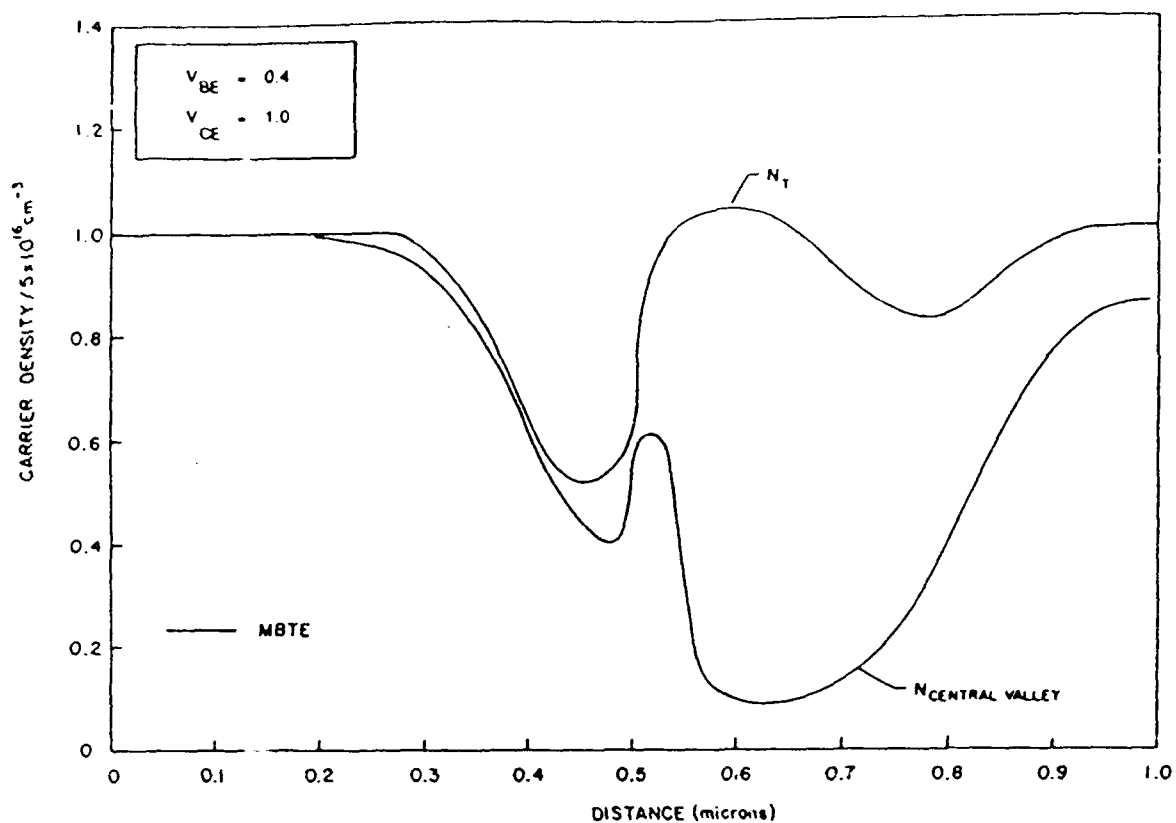Figure 2. PBT potential distribution along the central line of symmetry.

Figure 3. PBT carrier distribution along line of symmetry.

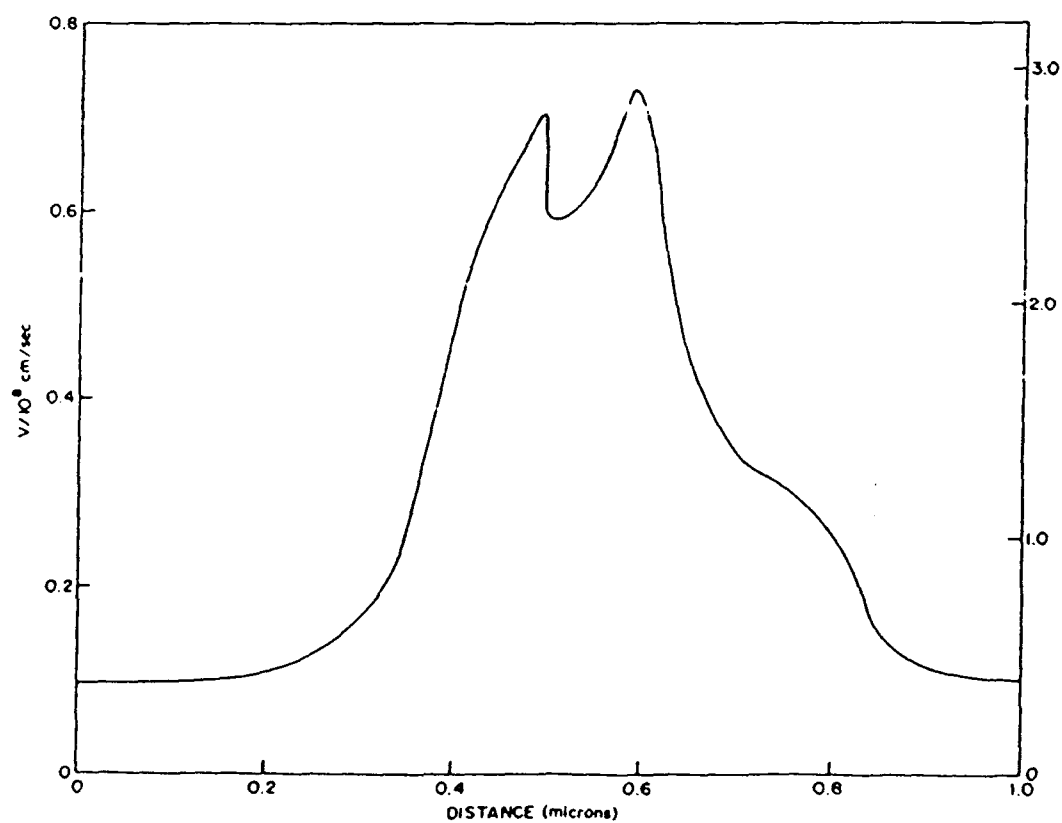Figure 4. PBT gamma valley velocity along line of symmetry.

F = 17.6 kV/cm

RISE TIME (ps)

100.0
50.0

10.0

5.0

1.0

0.5

0.1

2 3 4 5 6 7 8 9

(V x 10^{+7}) cm/sec

Figure 5. Dependence of peak overshoot
velocity on rise time. (From [3]).

CATHODE | N$^+$ | N$^-$ | N$^+$ | ANODE
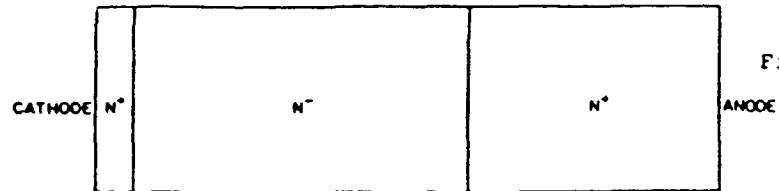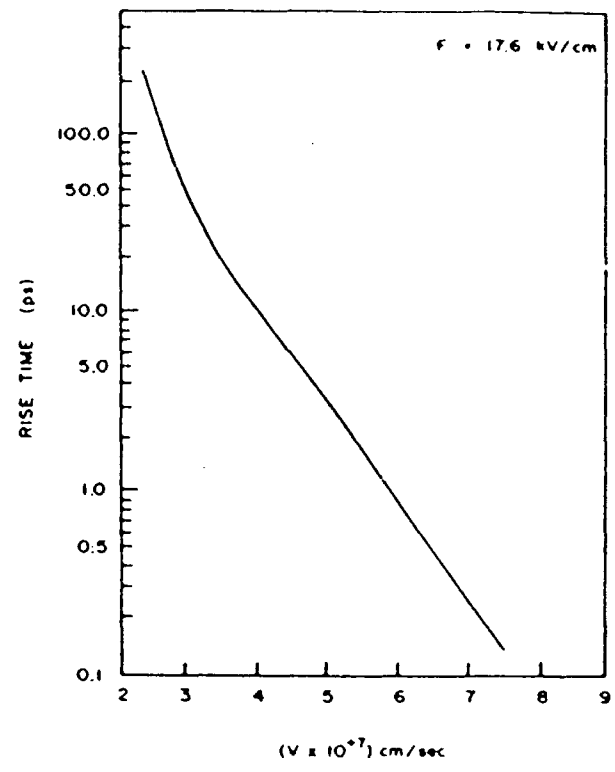
Figure 6. Schematic of structure for
examining displacement current
effects. (From [4]).

Figure 7. Initial current transient for
the structure of figure 6.

$\frac{J}{J_0}$

10.0

8.0

6.0

4.0

2.0

0.0

0.00 0.2 0.4 0.6 0.8 1.0 1.2 1.4

TIME (psec)
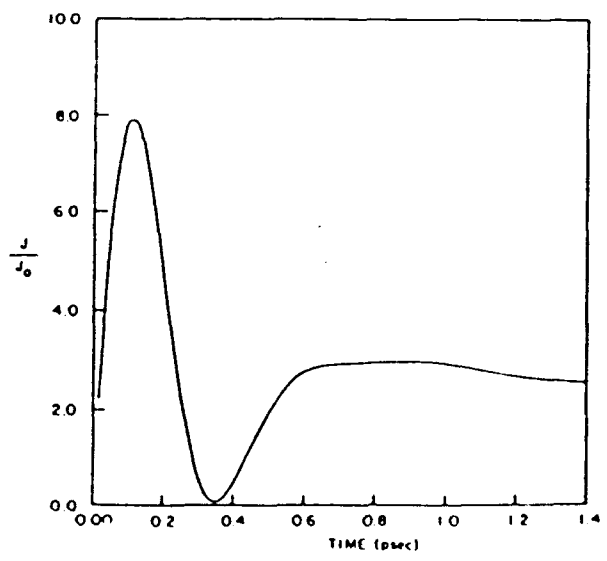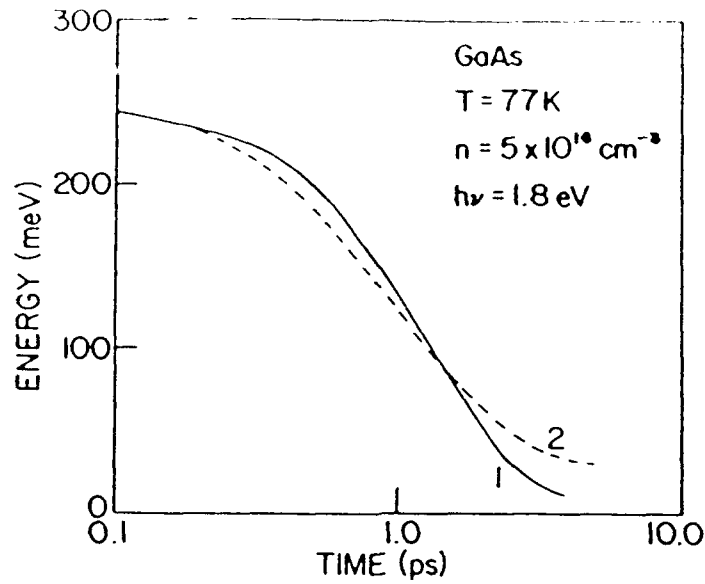
Figure 8. Energy relaxation with equilibrium
phonons (1) and nonequilibrium
phonons (2).
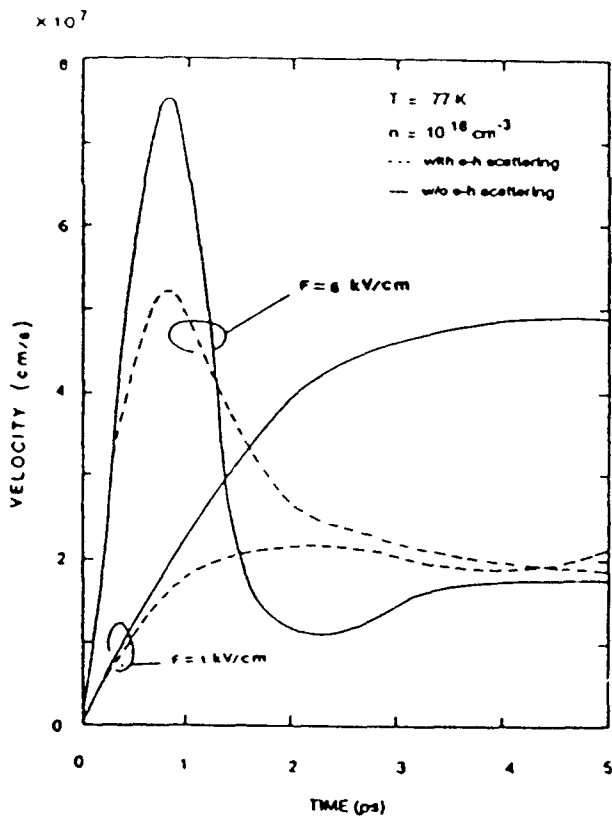
Figure 9. Overshoot with and without
e-h scattering. (From [5]).
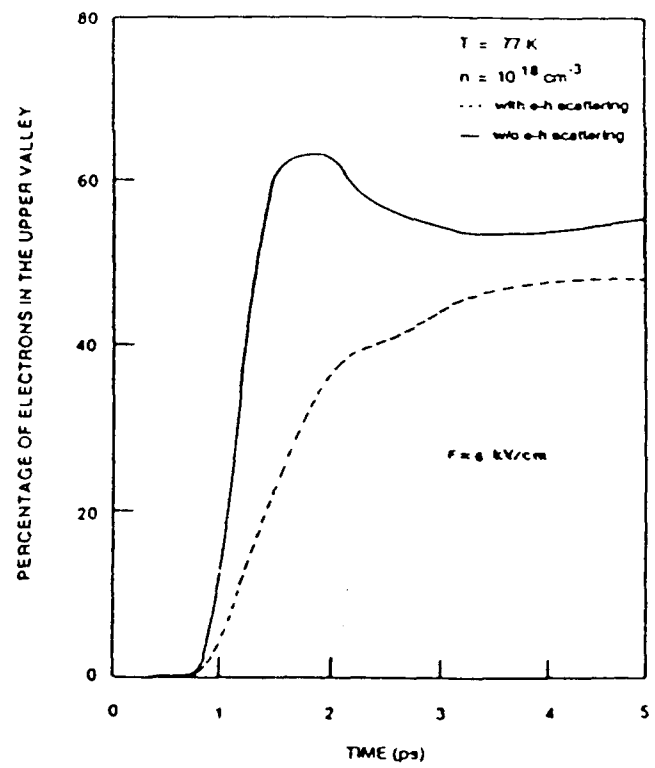
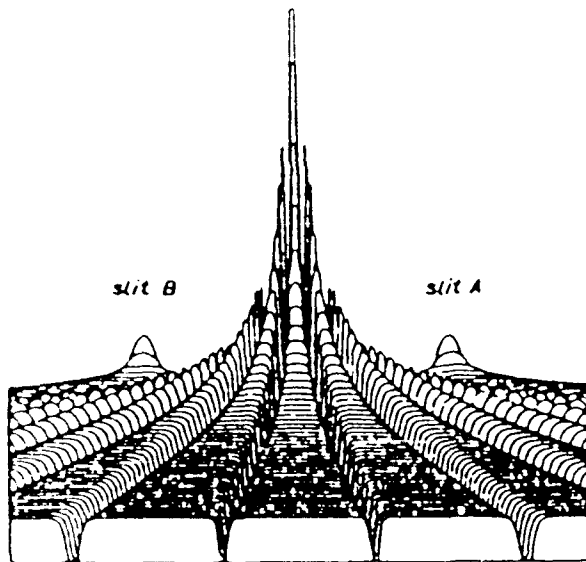Figure 10. Carrier distribution with and without
e-h scattering. (From [5]).

Figure 11. Quantum potential for the two-slit system. (From [8]).
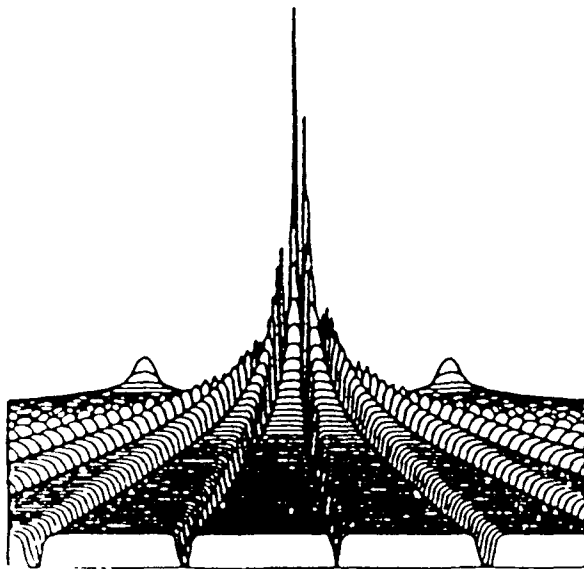

Figure 12. Particle trajectories for the two-slit system. (From [8]).

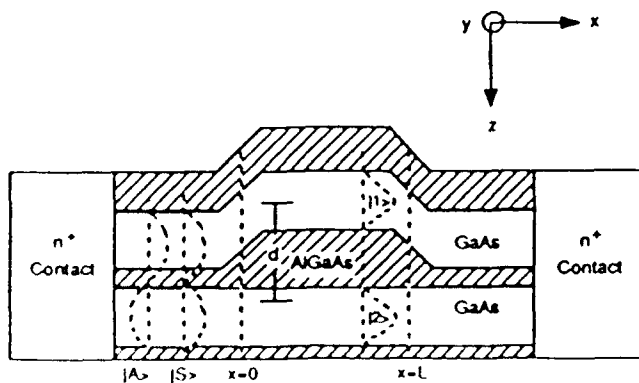Figure 13. Quantum potential for the AB effect (note the asymmetry). (From [8]).

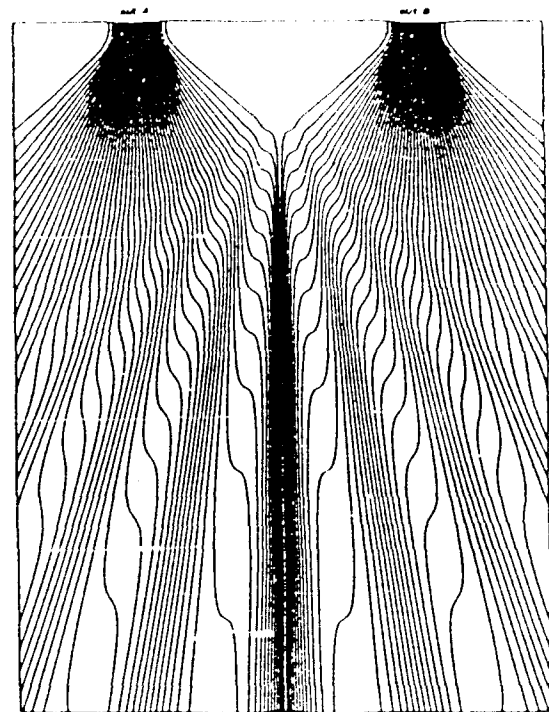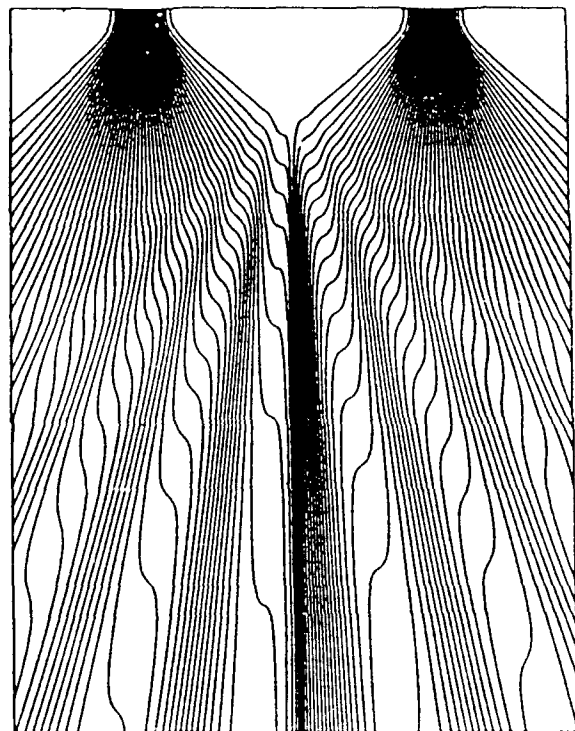Figure 15. Proposed structure for quantum interference effects. (From [9]).

Figure 14. Trajectories for the AB effect (note the shift of the overall pattern). (From [8]).